

The influence of study interests and (non-) cognitive predictors on study choice and study achievement in Flemish higher education.

Stijn Schelfhout

Supervisor: Prof. Dr. Wouter Duyck

Co-supervisor: Prof. Dr. Filip De Fruyt

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Doctor of Psychology

Academic year 2018–2019

THE INFLUENCE OF STUDY INTERESTS AND (NON-) COGNITIVE PREDICTORS
ON STUDY CHOICE AND STUDY ACHIEVEMENT
IN FLEMISH HIGHER EDUCATION.

Dedicated to the life and loving memory of Cyriel Schelfhout

24/01/1951 – 10/06/2015

Acknowledgements

This investigation on study interests came to fruition during a period I can only describe as trying times indeed, for both my family as myself. As such, this enterprise has not been a one man endeavor, but more of a strange and curious journey in which quite a few individuals have left wonderful social and artistic stamps on my passport in my search for the true nature of study interests. Although I do not have the space or time to thoroughly describe every realistic tree or conventional bush alongside the road I have traveled, I do feel that some of you are entitled to be mentioned by name and merits for the realization of this dissertation.

First and foremost, to my promotor, prof. dr. Wouter Duyck,

for putting his trust in me to let me make this dissertation my way and guiding me along the way. For also pointing out the obvious that the main effect could possibly be more important than the six-way interaction... and making me realize it... multiple times.

To my co-promotor, prof. dr. Filip De Fruyt,

for increasing my feeling with the field of vocational interests and for synchronizing my wordings with the sensitivities of editors and reviewers.

To prof. dr. Bart Wille,

for the countless reads and rereads of numerous manuscripts; stalking would not be exaggerated for the times I have severely shelled you with a bombardment of questions. For the encouraging words after yet another rejected manuscript.

To the other members of my guidance commission, prof. dr. Eva Deros, prof. dr. Filip Lievens and dr. Katja Pässler,

for the most welcome advice during these four years and for the permission to continue my research.

To my colleagues, dr. Lot Fonteyne and Miss Elizabeth Roels,

for listening to my sometimes weird ideas and giving me appropriate feedback when needed. It was my privilege of working with such talented individuals.

To Mrs. Lies Baten and Mr. Christophe Blanckaert,

for services rendered which are duly appreciated, regarding the ungodly art of making SAP-reports and the care for my digital soul. It is little things like these that makes my job a small piece of heaven on this earth.

To Miss Laure Dombrecht,

for extending her workings beyond our master years as beacon of advice and sometimes keeper of my (in)sanity.

To my Mother, Mrs. Ingrid Vanvaerenbergh,

for just being there, alongside me, all of my life. What would a researcher be without nightly copious meals, a counselor on standby for 24/7 and the unconditional support and belief in everything I ever attempted.

Whatever the outcome may be, for all this, and much more,

Thank you.

I will never forget this.

Stijn Schelfhout, June 2019

Table of Contents

Dissertation Abstract.....	9
Nederlandstalige Samenvatting	12
Chapter 1. Introduction.....	15
1.1. Dissertation Goal.....	16
1.2. The SIMON Project.....	18
1.3. Study Interests and Person – Environment Fit.....	20
1.4. The Effects of Study Interests on Study Choice.....	23
1.5. The Effects of Study Interests on Study Achievement	25
1.6. References.....	30
Chapter 2. From Interest Assessment to Study Orientation: an Empirical Advice Set Engine.....	38
2.1. Abstract	39
2.2. Introduction.....	40
2.3. Method and Materials.....	50
2.4. Results	60
2.5. Discussion.....	65
2.6. References.....	73
2.7. APPENDIX A: SIMON-I Questionnaire.....	77
2.8. APPENDIX B: EASE Manual and Executable RStudio Code: an Example for one Student.....	81
Chapter 3. The Influence of Interest Fit on (STEM) Study Choice: Female Students Fit their Choices Better.....	91
3.1. Abstract	92
3.2. Introduction.....	93
3.3. Method and Materials.....	99
3.4. Results	105
3.5. Discussion.....	110
3.6. References.....	116
3.7. APPENDIX A: SIMON-I Questionnaire.....	121
Chapter 4. Identifying Students at Risk of Failing in Academic Higher Education.	125
4.1. Abstract	126
4.2. Introduction.....	127
4.3. Method and Materials.....	136
4.4. Results	143
4.5. Discussion.....	146
4.6. References.....	151
4.7. APPENDIX A: Measures: Description, Reliability and Validation.....	158

Chapter 5. The Effects of Vocational Interest on Study Results: Student Person–Environment Fit and Program Interest Diversity.....	163
5.1. Abstract	164
5.2. Introduction.....	165
5.3. Method and Materials.....	174
5.4. Results	182
5.5. Discussion.....	194
5.6. References.....	202
5.7. APPENDIX A: SIMON-I Questionnaire.....	208
Chapter 6. General Discussion.	212
6.1. General Setting.....	213
6.2. Empirical Findings.....	214
6.3. Theoretical Impact of the Dissertation.....	219
6.4. Practical Impact of the Dissertation	222
6.5. Limitations and Future Research.....	223
6.6. Conclusion	224
6.7. References.....	225
Data Storage Fact Sheets.....	232
Reference List.....	248

Dissertation Abstract

The process of making the transition towards higher education is a trying experience for each Flemish student. As Flanders has an open access higher education, any student with a high school degree can choose to study nearly all available programs. To obtain the desired degree, a student has to successfully navigate two assignments. First, the student has to make an attainable study choice. And second, the student has to achieve study success at exams to stay on track for obtaining the degree. These assignments are not trivial. In this dissertation, I have presented data showing that only 36% of all students in the first year manages to pass all courses to stay on track for timely degree attainment. In order to improve this alarming number, Ghent University has started the SIMON-project (Study Skills and Interest MONitor) in an attempt to improve study success. SIMON aims to (re)orient students at risk to a more attainable program before they fail their exams and lose a lot of time and resources. Using a student's study interests depicted on the RIASEC (realistic, investigative, artistic, enterprising, conventional) hexagon, the current dissertation aims at improving the future impact of the SIMON project by investigating a number of pending issues in literature regarding study choice and study achievement. To this extent, I have presented results on four empirical studies. I have thus found that the PE (person-environment) interest fit of a student and a pool of study programs can be approximated through a uniform distribution. This distribution can be used as a baseline for the Empirical Advice Set Engine or EASE. EASE provides custom made study orientation based on an objective balance criterion. EASE thus dispenses advice sets of interesting programs for each individual student, while balancing the size and the minimal fit of these advice sets. Moreover, this EASE balance was superior to the balance rendered by the classic indices of PE interest fit, which are also used in SIMON. Also, such advice sets can be of capital importance towards study orientation for certain fields of expertise. Indeed, I have also investigated how the fit between the interests of students and

programs can contribute to an economically important STEM (science, technology, engineering, mathematics) choice. Results indicated that female students had a better fit with their program of choice (STEM and non-STEM) compared to male students. Female STEM students also had a better fit with the STEM field compared to their male colleagues. STEM study choice and the STEM gender gap towards male overrepresentation were both explained using a model featuring all six RIASEC dimensions, weekly hours of mathematics in high school and a fit with the STEM field. Male STEM study choice was more linked to the hours of mathematics in high school, while female STEM study choice was more linked to STEM field fit. Besides study choice, the present dissertation also deals with study achievement. As such, I have presented a change in methodology. This dissertation thus focuses on identifying students at risk through program-specific models, by predicting the achievement of individual students instead of explaining the population variance in study achievement, as is common in literature. The presented methodology also validates a number of cognitive and non-cognitive predictors towards identification of these students. For this identification, I have also investigated the possibility of using more lenient false positive rates (passing students identified as failing). Moreover, variables of study interests specifically, were present in 24% of the program-specific models, third only to study antecedents and cognitive ability. Regarding the influence of study interests on achievement, I have also investigated the role of the environment. Results indicated that programs showed a low diversity in study interests for their student populations. Student populations with higher diversity were also linked to higher average controlled motivation and lower average autonomous motivation. However, this low diversity did have enough variance to influence study achievement. In general, programs with higher diversity showed better study achievement. This effect reversed in programs with a specific RIASEC profile, high on the social dimension and low on the realistic dimension. I also found that the interest diversity of programs had a much more profound influence on

study achievement than individual student PE interest fit. To conclude, the dissertation's empirical results and specific operationalization of the RIASEC dimensions and PE interest fit provide a unique, open access focus on the effects of study interests. Indeed, orientation towards an interesting study choice is now based on an objective criterion of how well a student's study interests have to fit possible programs. Moreover, orientation towards attainable programs is now based on identifying students at risk by focusing on predicting individual study achievement, while still accounting for the specific, open access set up of the education system. This dissertation also enables counselors to immediately translate these findings towards practical application.

Nederlandstalige Samenvatting

De invloed van studie-interesses en (niet-) cognitieve predictoren op studiekeuze en studiesucces in het Vlaams hoger onderwijs.

De overgang naar hoger onderwijs is een beproeving voor elke Vlaamse student. Inderdaad, omdat Vlaanderen een open toegang heeft tot hoger onderwijs, kan elke student met een diploma secundair onderwijs aan bijna elke opleiding beginnen. Om het vooropgestelde diploma te halen van de gekozen opleiding, dient een student twee taken tot een goed einde te brengen. De student dient een haalbare studiekeuze te maken. En de student dient te slagen in examens om op schema te blijven om het begeerde diploma te halen. Deze taken zijn niet zo eenvoudig als ze lijken. De data in deze dissertatie geven aan dat slechts 36% van de eerstejaarsstudenten erin slaagt om alle opleidingsonderdelen tot een goed einde te brengen om zo op schema te blijven om tijdig het beoogde diploma te behalen. Om dit onrustwekkende cijfer te verbeteren heeft de Universiteit Gent het SIMON-project (Study Skills and Interest MONitor) gestart. SIMON is erop gericht om studenten die dreigen te falen te (her)oriënteren naar een meer haalbaar studieprogramma vooraleer ze hun examens effectief falen, met verlies van tijd en middelen tot gevolg. In deze dissertatie worden de PAKSOC (praktisch, analytisch, kunstzinnig, sociaal, ondernemend en conventioneel) studie-interesses van een student gebruikt om de impact van het SIMON project te vergroten door in de literatuur een aantal openstaande vragen te onderzoeken omtrent studiekeuze en studiesucces. Om dit te bewerkstelligen heb ik in deze dissertatie de uitvoering en resultaten besproken van vier empirische studies. Zo heb ik onder meer gevonden dat de fit tussen de interesses van een student en een set van studieprogramma's kan worden benaderd via een uniforme distributie. Deze distributie kan dan worden gebruikt als de basis voor een Empirische Advies Set Engine, of ook wel EASE. EASE verstrekt gepersonaliseerde studieoriëntatie voor elke student, gebaseerd op een objectief criterium dat toelaat de lengte

en de fit van de set met voorgestelde programma's te balanceren. Deze balans is superieur aan deze die wordt gegenereerd door meer klassieke indices van interessefit, die trouwens ook worden gebruikt in SIMON. Dergelijke studieoriëntatie kan van cruciaal belang zijn in bepaalde gespecialiseerde gebieden. Als dusdanig heb ik ook onderzocht hoe de interessefit van studenten en studieprogramma's kan bijdragen tot een economisch belangrijke STEM (wetenschap, technologie, ingenieur en wiskunde) studiekeuze. De resultaten hiervan laten duidelijk zien dat vrouwelijke studenten een betere interessefit hadden met hun gekozen (STEM en niet-STEM) programma in vergelijking met mannelijke studenten. Vrouwelijke STEM - studenten hadden ook een betere interessefit met het STEM veld in vergelijking met hun mannelijke collega's. STEM studiekeuze en de genderkloof (mannelijke meerderheid) in het STEM veld werden verklaard door een model dat alle PAKSOC dimensies bevatte, naast wekelijkse uren wiskunde in het secundair onderwijs, en de fit met het STEM veld. Een mannelijke STEM keuze was meer gerelateerd aan uren wiskunde in het secundair, terwijl een vrouwelijke STEM keuze meer gerelateerd was aan de fit met het STEM veld. Naast studiekeuze behandelt de huidige dissertatie ook studiesucces. Omtrent dit studiesucces, heb ik ook een verandering voorgesteld in methodologie. Als dusdanig spitst de huidige dissertatie zich toe op identificeren van studenten die dreigen te falen in hun gekozen studieprogramma. Hiertoe heb ik mij vooral gericht op het voorspellen van resultaten van individuele studenten, en niet op het verklaren van populatievariatie in studiesucces, zoals het meestal gebeurt in de literatuur. Deze methodologie valideert ook een set (niet-) cognitieve predictoren voor identificatie van falende studenten. Wat betreft deze identificatie, heb ik ook de mogelijkheid onderzocht om minder strenge vals-positieve (succesvolle studenten die worden geïdentificeerd als falend) ratio's te gebruiken. Specifiek voor studie-interesses heb ik een aanwezigheidsgraad gevonden van 24% in de identificatiemodellen. Dit betekent dat studie-interesses voorkwamen in 24% van de (programma-) specifieke modellen om studiesucces te

voorspellen. Dit was de derde meest impactvolle predictor, na studieantecedenten en cognitief vermogen. De relatie tussen studie-interesses en studiesucces wordt ook beïnvloed door de omgeving. Resultaten laten zien dat programma's een lage diversiteit hebben in de studie-interesses van studenten die het programma hebben gekozen. Populaties met een hogere diversiteit werden trouwens gelinkt aan hogere gemiddelde gecontroleerde motivatie en lagere gemiddelde autonome motivatie. In het algemeen was een hogere diversiteit over programma's ook gelinkt aan betere gemiddelde studieresultaten. Bij een aantal programma's met een zeer specifiek interessepatroon (hoge sociale dimensie, lage praktische dimensie) observeerde ik echter het omgekeerde effect. Ik vond ook dat de interessediversiteit in programma's een sterkere invloed had op studiesucces dan individuele interessefit. Om te besluiten, stel ik dat in deze dissertatie, de empirische resultaten en de specifieke operationalisatie van de PAKSOC dimensies en interessefit een uniek perspectief (open toegang) bieden op studie-interesses en hun effect op studiekeuze en studiesucces. Oriëntatie naar een interessante studiekeuze wordt gebaseerd op een objectief criterium: hoe goed moet de fit zijn tussen de interesses van een student en het profiel van een programma? Oriëntatie naar haalbare studiekeuzes wordt gebaseerd op het identificeren van falende studenten door het voorspellen van studiesucces, terwijl er nog altijd wordt rekening gehouden met de specifieke set up van het onderwijssysteem met open toegang. Deze dissertatie stelt studieadviseurs ook in staat deze bevindingen onmiddellijk in de praktijk te brengen.

Chapter 1. Introduction

1.1. Dissertation Goal

The transition from high school to higher education is a challenging process all (Flemish) students have to face on their way towards attaining the desired degree (Tett, Cree, & Christie, 2017). Moreover, due to its open access and low cost system, Flemish higher education has a specific higher education set up (Eurydice, 2017; Organization for Economic Co-operation and Development or OECD, 2017a). Indeed, barring a few exceptions like medicine, dentistry and performance arts, all study programs are open to anyone with a high school degree for a modest yearly tuition fee, not including possible scholarships (for a complete overview see *Flemish Education in Figures, 2016-2017*). As described by Fonteyne (2017), Flemish higher education thus features two key assignments students have to navigate successfully in order to attain their degree. First, students have to make a well thought out study choice towards a study program. And second, students have to achieve study success at exams to stay on the model trajectory of their study program. These two assignments towards degree attainment are not so straightforward as they might seem. As an example, recent Ghent University data from 2016 to 2018 (featuring in this dissertation) indicate that first year study success across bachelor students is quite low, with only 36% of all students taking exams staying on track to get their degree in timely fashion (model trajectory + one year). In contrast, 52% of the students does not stay on track of their model trajectory, while only 12% reorients or drops out before taking exams. In *The State of Higher Education 2015-2016*, the OECD (2017b) reports that 30% of the students leaves higher education without degree in the OECD countries (Canada, Europe and the United States). This report clearly shows that the low success rate is not a problem exclusively to Flanders. The 2015-2016 report (OECD, 2017b) also points out that not everyone who starts higher education has the same basic skillset to begin with. This implies that some students lack the basic skills to succeed in their program of choice.

In order to improve the success rate of students, Ghent University already started the SIMON (Study capacities and Interest MONitor) project (Fonteyne, 2017) to (re)orient more students at risk towards a suitable program, based on their skills and interests. Indeed, timely reorientation saves a lot of resources (up to € 12,000 for each student see EUROSTAT, 2017) and time (at least one year) for both students and higher education institutions (OECD, 2017b). To this end, SIMON specifically focuses on identifying those students who do not have the basic skills to succeed in their program of choice. First results indicate that SIMON indeed manages to identify more than 13% of all students at risk (Fonteyne, Duyck & De Fruyt, 2017), while also making an impact on reorientation through (negative) feedback (Fonteyne et al., 2018). Though the SIMON project is already working as intended, SIMON's future impact can be further improved by addressing a number of pending issues in literature regarding study choice and study achievement.

The present dissertation wants to address these issues in literature and in the SIMON project by taking the study interests of students as primary focus. These study interests reflect the vocational interest(s) of students prior to and during their study program. In other words, which activities or professions do students consider interesting enough to personally do or pursue? Literature already indicates that students' study interests are predictive of both study choice (Burns, 2014; Donnay, 1997; Kuder, 1977; Päßler & Hell, 2012; Rounds & Su, 2014; Whitney, 1969) and study achievement (Nye, Butt, Bradburn, & Prasad, 2018; Nye, Su, Rounds, & Drasgow, 2012; Rounds & Su, 2014; Tracey, Allen, & Robbins, 2012).

This dissertation is built around an introduction (present Chapter 1), four empirical chapters (Chapters 2 to 5) and a discussion (Chapter 6). In Chapter 2 of this dissertation I have investigated which objective criterion should be used when advising specific programs to an individual student. In Chapter 3, I have investigated how the fit between the interests of students and study programs contributes to an economically important STEM (science,

technology, engineering, mathematics) study choice (Perera & McIlveen, 2018) and to an explanation of the still expanding gender gap in STEM enrolment (Stoet & Geary, 2018; Xu, 2008). In Chapter 4, I have investigated how well we can identify failing students in academic higher education through the use of study interests and other cognitive and non-cognitive variables. And finally, in Chapter 5, I have investigated the effects of the environment on the relationship between the interests of students and study results. The ultimate goal of this dissertation is twofold. First, it aims to add to literature regarding vocational interests and their influence on study choice and study achievement. And second, it also aims at improving the SIMON project towards timely degree attainment. For this purpose, in Chapter 6, the results regarding the four presented empirical studies are integrated into the current literature and the practical consequences (for the SIMON project) are discussed.

1.2. The SIMON Project.

The data for this dissertation were obtained within the scope of the SIMON project (Fonteyne, 2017). SIMON's ultimate goal is to improve timely degree attainment by providing study (re)orientation advice to students prior to and during the initial phases of higher education by matching them to attainable and interesting programs. The project originated from a mere mathematics test to establish the entry levels of a student's mathematical knowledge at the start of a program at the Faculty of Psychology and Educational Sciences (Fonteyne et al., 2015). Interestingly, the test seemed to have predictive validity for passing the statistics exam and for passing the first year of the program altogether. Indeed, results showed that the test could explain more than 4% of the variance of student success rates in the first year. Though this number seemed quite low, it did give an indication that some students do not have the basic (cognitive) skills to succeed in their chosen program and that these students could be identified. By constructing a full test battery of known predictors of study achievement, students at risk of failing could be identified to upgrade their

basic skills or reorient them towards a more suitable program altogether, months before they would fail their exams. This full battery of tests (called SIMON-S) to predict study achievement was constructed shortly thereafter, containing predictors regarding study background, cognitive ability, motivation, metacognition, academic self-efficacy and test anxiety (Fonteyne, Duyck, et al., 2017). The instrument is used to orient students towards attainable programs. Alongside SIMON-S, another instrument was devised to give additional orientation information to future students based on their study interests. This SIMON-I interests questionnaire provided interest profiles of which more than 91% of the respondents agreed it reflected their interests well enough (Fonteyne, Wille, Duyck, & De Fruyt, 2017). As it stands, the project now features three data streams that are used to improve the instrument and facilitate scientific research. The first stream originates from within Ghent University and presents SIMON-S to all first year students to validate the instrument. The test is not mandatory, but students are strongly encouraged to participate, which leads to response rates that are rather high (an average response rate of about 68% for the data ranging from 2016 to 2018). These (historical) data are then linked to exam results to establish or fine tune study achievement models. As such, these models enable the prediction of study achievement for (future) students. Students themselves also receive feedback on their test scores. The first series of results indicate that the (sometimes negative) feedback does indeed affect reorientation (Fonteyne et al., 2018). The second stream also originates from within Ghent university and presents SIMON-I to third year bachelor or first year master students who have achieved study success and persistence in their program of choice. Their study interest profiles are used to establish interest profiles for their chosen study programs. The third and final stream is the online application of SIMON-S and SIMON-I and can be consulted at www.vraaghtaansimon.be by future students making the transition towards higher education. After taking both tests, the future students receive an overview of which programs could

prove interesting based on their study interests and which programs are attainable based on prediction models of (non-) cognitive predictors. Though this dissertation makes use of all data streams within the project, nearly all data used (with exception of the construction of program interest profiles as the program profiles benefit from larger datasets) are independent of previous publications originating from within the project (Fonteyne, 2017; Fonteyne et al., 2015; Fonteyne, Duyck et al., 2017; Fonteyne et al., 2018; Fonteyne, Wille et al., 2017).

1.3. Study Interests and Person – Environment Fit

Rounds and Su (2014) describe vocational interests as traits that drive people towards certain activities, towards the environment in which these activities occur or towards the outcomes of these activities. For the present dissertation on study interests, these activities specifically comprise of studying particular topics, in the environment of a specific study program, with the outcome of obtaining the specific degree. Derived from this description, Rounds and Su (2014) also highlighted a number of properties of vocational interests that are crucial to this dissertation and its focus on study interests. Indeed, vocational interests are (1) stable traits indicating individual differences that are (2) contextualized, while also giving a (3) direction to behavior (study choice), with the result of being (4) predictive of behavior (study achievement).

First and foremost, vocational interests are stable traits that exhibit individual differences (Wille, De Fruyt, & Feys, 2010). Since the fifties of previous century, the leading model describing these differences incorporates six RIASEC (realistic, investigative, artistic, social, enterprising and conventional) dimensions of vocational interests (Holland, 1963; Holland, 1997). Moreover, these dimensions are organized in a clockwise hexagonal fashion that has been empirically verified (Tracey, 1997; Tracey & Rounds, 1996). The model is still used in research today due to its core tenet of commensurate measurement of persons and environments (Tracey et al., 2012). In other words, the study interests of students and their

chosen programs can be depicted on the same RIASEC dimensions. As such, I have used these dimensions throughout this dissertation to describe the study interests of students (persons) and their programs of choice (environments). Moreover, as study interests remain quite stable throughout the course of higher education (Swanson & Hansen, 1988), it also becomes possible to investigate their influence on study choice and study achievements in prediction or prospective studies.

To facilitate this description of study interests, I have made use of a second property of vocational interests as described by Rounds and Su (2014). Indeed, interests are always contextualized, they always have an object like an activity, an environment or a goal. This is another crucial property as it enables the construction of interest questionnaires. Responses to such a questionnaire can then be used to profile (using the RIASEC dimensions) the study interests of students and programs. Literature already features a vast amount of instruments that determine a person's RIASEC profile by providing a score on each of the dimensions (Arbona, 2000; Rayman & Atanasoff, 1999). This dissertation however, uses the SIMON-I questionnaire as it was specifically developed to provide custom made orientation advice for students making the transition towards higher education (Fonteyne, Wille et al., 2017). The instrument consists of 153 short items. Students have to indicate on a yes or no scale which activities or professions they find interesting. Examples include “*forrester*” (R) or “*writing a scientific paper*” (I) which load on one of the six RIASEC scales. Depending on the answers, a RIASEC profile is established for each student that contains a score (0 – 100) for each dimension. In contrast, to profile a program, I have opted to use the incumbent method proposed by Allen and Robbins (2010). As such, the RIASEC dimension scores for a specific study program profile are obtained by averaging out the RIASEC scores of successful and persistent students in that specific program, which are drawn from the second data stream of the SIMON project (see above).

As both student and study program profiles can be established using the same RIASEC dimensions, it now becomes possible to make the person-environment interest fit (PE interest fit) between student and study program. Literature already features two groups of methods how PE interest fit can be operationalized in research. First, classic approaches to PE interest fit rely on high point coding to obtain a congruence index (Brown & Gore, 1994; Young, Tokar, & Subich, 1998). In such an index, the RIASEC dimension scores for students and study programs are ranked from high to low. By comparing the rank and placement of the letters in the student and program RIASEC profiles, a non-continuous (ordinal at best) measure of interest fit is established. Although these classic congruence indices are transparent and user friendly, they also have limitations (Tinsley, 2000; Tracey & Robbins, 2006). To give one example, most classic methods do not use the entire RIASEC profile to establish a PE interest fit, in contrast to the plea of Holland (1997) himself to always take into account the full profile. In reaction to such limitations, continuous measures of PE fit have emerged using all RIASEC dimensions. The current dissertation features two of these methods as they show criterion validity towards study achievement in the first (correlation fit) and third year (Euclidean distance) of higher education (Tracey et al., 2012). Correlation fit is calculated by making the mere correlation between the scores on the RIASEC profiles of student and study program. Euclidean distance calculates the distance between RIASEC profiles of student and study program on a two-dimensional (people / things and data / ideas) overlay of the RIASEC hexagon (Prediger, 1982).

Rounds and Su (2014) consider the direction towards a (study) choice and the prediction of (study) behavior an important third and fourth property of vocational interests. As such, I have made a brief summary of both topics in the next two sections, describing the issues addressed in the empirical chapters (Chapters 2 to 5) of this dissertation.

1.4. The Effects of Study Interests on Study Choice

Despite the new continuous methods, the use of PE interest fit towards study orientation is still faced with a problem as there is no objective criterion of what exactly determines a good fit between a student and a study program. To this date, as there are hardly any alternatives, literature and practice still resort to obsolete heuristics to match student and programs, based on common practice and not on theoretical or empirical evidence. As a consequence, literature also does not provide an answer to how many study programs counselors should advise to students as part of an advice set towards higher education. As a practical example, within the SIMON project, the third data stream provides orientation advice to students by combining the *one-letter agreement index* (Holland, 1963) and the *two-letter agreement index* (Healy & Mourton, 1983). The former index defines a good fit as a match between the highest dimension of student (RIASEC) and program (RAISEC) profiles. The latter index defines a good fit as a match between the two highest dimensions of student (RIASEC) and program (IRASEC) profiles. Within SIMON, a good fit is thus defined as a one-letter agreement or a two-letter agreement between two of the highest three dimensions of the student profile and the highest two dimensions of the study program profile. As such, RAISEC (student) and RSECAI (program) is a good fit, as well as RAISEC and IASECR, but not RAISEC and ASRECI. However, such a relative criterion for a good fit is only valid within a specific setting or instrument and does not have an objective or empirical criterion (Camp & Chartrand, 1992). To address this issue in Chapter 2, I present a novel Empirical Advice Set Engine (EASE) using the existing correlation method to make the fit between a set of programs and a student's study interests. As PE interest fit seems to be uniformly distributed within a person, I have used this uniform distribution as the basis for an Empirical Advice Set Engine or EASE, by balancing the number of presented study options in an advice set against the PE interest fit of the options in the advice set. As such, EASE objectively

establishes an individual threshold how well each student needs to fit a program before the program can be considered interesting enough to present it to the student as a valid study choice. The EASE methodology is validated by rendering advice sets for each individual student and by comparing these advice sets to the advice sets generated by the classic methods used in SIMON.

Such custom made advice sets can prove to be crucial towards getting sufficient enrolments and subsequent influx into important work fields. Recently, industrialized regions have experienced a decline in STEM (science, technology, engineering and mathematics) enrolments towards higher education (Ainley, Kos, & Nicholas, 2008; Perera & McIlveen, 2018). This decline can have huge ramifications as filling out vacancies in the STEM field is key to maintaining the economic success of these regions (World Economic Forum, 2016). Apart from the decline in enrolments, literature also indicates that there is a gender gap in STEM enrolments towards female underrepresentation (Xu, 2008). A recent study even indicates that this gap is widening, especially in industrialized and gender-aware countries (Stoet & Geary, 2018). For sure, the properties of the STEM field and the gender gap are already highlighted in vocational interests literature. Meta-analytic evidence from within the STEM field shows that female students score higher on the social dimension, while male students have higher scores on the realistic and investigative dimension (Su, Rounds, & Armstrong, 2009). Also, different explanations have been formulated to explain the gender gap including variables like (early) exposure to science or performance on high school mathematics (Dejarnette, 2012; Wang, 2013a; Wang, 2013b). Moreover, the representation of gender in different professions in the STEM field can be largely explained by the position of those professions on a people / things dimension overlay (Su & Rounds, 2015). However, the research on the decline in enrolments still shows major blind spots. First, the differences in study interests between students choosing a STEM career and students who do not, were

never compared against each other. Second, PE interest fit has never been used to explain a (non-) STEM study choice. And third, literature is also oblivious if the student's gender could influence the PE interest fit with the program of choice, within and even outside the STEM field. Such a difference could allow us to enlarge the knowledge on the gender gap and how to close it. To address this issue in Chapter 4, I have investigated if there was a difference between male and female students regarding PE interest fit with their (STEM and non-STEM) programs and the STEM field. Moreover, I have also investigated how this PE interest fit would contribute to a prediction of STEM study choice and an explanation of the gender gap above and beyond the RIASEC dimensions and hours of high school mathematics.

1.5. The Effects of Study Interests on Study Achievement

The effect of study interests on study achievement is normally assessed through PE interest fit. In other words, does the interest fit between a student and the chosen program predict how this student will perform in higher education? Although literature already agrees that different programs will reward different student interest RIASEC patterns (Smart, Feldman, & Ethington, 2000), the debate regarding the size of the effect is still ongoing. Meta-analytic evidence on the effects of PE interest fit on study achievement range from marginal at best, to effects of around $r = .32$ or about 10% explained variance, while also displaying incremental validity above and beyond the mere height of single RIASEC dimension scores (Nye et al., 2012). These effects are not nearly as high as Holland predicted they would be (Holland, 1997). As an explanation, several sources suggest the discrepancy between theoretical expectations and empirical results could be caused by reducing the six dimensional RIASEC construct to a one-dimensional fit. A recent string of research using polynomials has reported a promising new approach (Nye, Prasad, Bradburn, & Elizondo, 2018). By making the fit in function of the purpose it has to serve (predicting study achievement), the power of detecting effects of PE interest fit on study achievement should

become larger. For instance, by explaining GPA (grade point average) through a combination of all RIASEC dimensions and their interactions from both student and environment profiles (rendering a polynomial of 36 terms), results show that the method clearly trumps classic high point coding congruence indices regarding the effect of PE interest fit. Indeed, using this method, PE interest fit can explain about 10% of the population variance in GPA, but more importantly also displays about 3% incremental predictive validity on top of models including cognitive ability, situational judgement and biodata (Nye, Bradburn et al., 2018). The method also has one downside as it always needs a (continuous) dependent variable as a target, making it unwieldy to use in a context of study choice.

As the basis for possible (re)orientation advice, the SIMON project tries to predict study achievement in first year students by making use of the first data stream. By comparing test results on SIMON-S to historical data in program-specific models, a prediction can be made about study success on exams, and whether or not the student will pass (Fonteyne, Duyck et al., 2017). For (re)orientation advice, SIMON specifically targets those students who do not have the basic skillset to succeed. SIMON only gives a negative advice if the model is 95% sure the student will fail. Based on meta-analytic evidence, the program-specific regression models incorporate a pool of known predictors of academic achievement that explain various amounts of population variance. For a complete overview and discussion on these predictors, I refer to meta-analytic evidence on the predictors of study achievement in general (Schneider & Preckel, 2017), and to specific meta-analytic research and studies on study antecedents (Hodara & Lewis, 2017; Poole, Shulruf, Rudland, & Wilkinson, 2012), cognitive ability (Rohde & Thompson, 2007; Roth et al., 2015), motivation (Kriegbaum, Becker, & Spinath, 2018; Ryan & Deci, 2000; Ryan & Deci, 2017), personality (De Fruyt & Mervielde, 1996; Furnham, Chamorro-Premuzic, & McDougall, 2003), academic self-efficacy (Bandura, 1993), test-anxiety (Credé & Kuncel, 2008), self-control (Duckworth,

Taxer, Eskreis-Winkler, Galla, & Gross, 2019) and metacognition (Kitsantas, Winsler, & Huie, 2008). Results from SIMON (up until 2016) show that study antecedents (7%), cognitive ability (8%), motivation and test anxiety (7%), personality - conscientiousness (3%), all explain a wide range of incremental population variance in the study achievement (GPA) of first year students (Fonteyne, Duyck et al., 2017). Program-specific models containing these predictors were able to explain an average of about 23% in study achievement, while also identifying more than 13% of the students at risk of failing, with a maximum false positive rate of 5% (passing students identified as failing).

In Chapter 4, I have presented two alterations to the SIMON methodology as it also features in literature (Fonteyne, Duyck et al., 2017). First, I have included the students' study interests in the pool of possible predictors of study achievement. To this extent, I have modified the polynomial fit approach (Nye, Bradburn et al., 2018; Nye, Prasad et al., 2018). Indeed, as the SIMON project uses program-specific models, the RIASEC environments within a program-specific model remains stable, making all environment terms within the polynomial redundant as they remain constant. I have thus reduced the polynomial for each student to twelve terms only, consisting of the student's RIASEC scale scores and their quadrats. Instead of regressing them separately on GPA to acquire the polynomial fit, I have added them directly to the possible pool of study achievement predictors. However, as it remains possible that these dimensions can interact with each other, I have also included both correlation fit and Euclidean distance as they represent an interaction of all six dimensions.

And second, I have also altered the focus of the SIMON methodology in Chapter 4. Though the SIMON project has already rendered promising results in predicting study achievement (Fonteyne, Duyck et al., 2017), the methodology used suffers from a major conceptual problem that runs rampant throughout psychological and educational literature. Indeed, models that are able to explain a lot of population variance do not necessarily predict

individual student results (Shmueli, 2010). Both uses of statistical modeling have been heavily conflated towards the use of an explanation of population variance for nearly all research questions. Within the SIMON project, this problem is especially relevant, as SIMON aims at identifying precisely those students who lack the basic properties to succeed in their program of choice. In Chapter 4 of this dissertation, I have thus presented a solution to this issue by using a program-specific methodology to predict individual student results and thus identify students at risk of failing. Instead of individual predictors, a cross-validated AIC (Akaike's Information Criterion) based methodology selects program-specific models that render the smallest prediction error towards study achievement (Burnham & Anderson, 2002). As such, Chapter 4 also validates a set of predictors towards prediction of individual student results, alongside the already documented explanation of population variance. To conclude this chapter, I have also explored the effects of using more lenient false positive rates (10%, 15% and 20%) on the identification rate of failing students.

Although the debate of the influence of study interests and PE interest fit on study achievement is very much alive, there is one very important element that has remained vastly understudied. Indeed, Nauta (2010) correctly pointed out that in vocational interest research, the environment did not receive nearly as much attention as the person, although both elements are equally important towards interest fit. For instance, literature still remains oblivious to the question how well students fit their program of choice. For sure, there are theories regarding homogeneity that predict that students with a similar profile in vocational interests should flock together towards similar study programs (Schneider, 1987). However, in practice these theories are hard to investigate, because the access to a lot of international education systems is locked behind entry exams or GPA requirements. Fortunately, higher education in Flanders is (almost) entirely open to anyone with a high school degree. As a result, such an environment is ideal to investigate the unbiased influence of vocational

interests on the homogeneity of study program populations. In Chapter 5 of this dissertation, I have thus investigated how diverse student interests are within and between study programs and how we can explain this diversity. Moreover, I have also investigated how diversity of study programs affects average study results in these programs. And finally, I have studied how these effects of program interest diversity compare to the effects of individual PE interest fit regarding study achievement.

1.6. References

- Ainley, J., Kos, J., & Nicholas, M. (2008). Participation in Science, Mathematics and Technology in Australian Education. *ACER Research Monographs*. Retrieved 11th of June 2019, from http://research.acer.edu.au/acer_monographs/4
- Allen, J. & Robbins, S. (2010). Effects of Interest–Study program Congruence, Motivation, and Academic Performance on Timely Degree Attainment. *Journal of Counseling Psychology*, 57(1), 23-35. doi:10.1037/a0017267
- Arbona, C. (2000). Practice and research in career counseling and development – 1999. *The Career Development Quarterly*, 49(2), 98–134. doi:10.1002/j.2161-0045.2000.tb00554.x
- Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist* 28(2), 117-148. doi:10.1207/s15326985ep2802_3
- Brown, S. D., & Gore, P. A. Jr. (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
- Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A practical information-theoretic approach* (2nd ed.), Springer-Verlag.
- Burns, S. T. (2014). Validity of Person Matching in Vocational Interest Inventories. *The Career Development Quarterly*, 62(2), 114-127. doi: 10.1002/j.2161-0045.2014.00074.x
- Camp, C. C., & Chartrand, J. M. (1992). A comparison and evaluation of interest congruence indices. *Journal of Vocational Behavior*, 41(2), 162–182. doi:10.1016/0001-8791(92)90018-U

- 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
- De Fruyt, F., & Mervielde, Y. (1996). Personality and interests as predictors of educational streaming and achievement. *European Journal of Personality*, 10(5), 405-425. doi: 10.1002/(SICI)1099-0984(199612)10:5<405::AID-PER255>3.0.CO;2-M
- Dejarnette, N. (2012). America's children: providing early exposure to STEM (Science, Technology, Engineering and Math) initiatives. *Education*, 133(1), 77–84.
- Donnay, D. (1997). E. K. Strong's legacy and beyond: 70 years of the Strong Interest Inventory. *The Career Development Quarterly*, 46(1), 2–22. doi: 10.1002/j.2161-0045.1997.tb00688.x
- Duckworth, A. L., Taxer, J. L., Eskreis-Winkler, L., Galla, B. M., & Gross, J. J. (2019). Self-control and Academic Achievement. *Annual Review of Psychology*, 70, 373-393. doi: 10.1146/annurev-psych-010418-103230
- EUROSTAT. (2017). Annual expenditure on public and private educational institutions per pupil/student. ec.europa.eu/eurostat/tgm/table.do?tab=table&plugin=1&language=en&pcode=tps00067
- Eurydice (2017). Retrieved November, 27, 2017, from <https://webgate.ec.europa.eu/fpfis/mwikis/eurydice/index.php/Countries>
- Flemish Education in Figures, 2016-2017, retrieved, June 6th 2019 from <https://www.vlaanderen.be/en/publications/detail/flemish-education-in-figures-2016-2017>

- Fonteyne, L. (2017). *Constructing SIMON : a tool for evaluating personal interests and capacities to choose a post-secondary major that maximally suits the potential*. Ghent University. Faculty of Psychology and Educational Sciences.
- Fonteyne, L., De Fruyt, F., Dewulf, N., Duyck, W., Erauw, K., Goeminne, K., Lammertyn, J., et al. (2015). Basic mathematics test predicts statistics achievement and overall first year academic success. *European Journal of Psychology of Education, 30*(1), 95–118. doi: 10.1007/s10212-014-0230-9
- 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. doi: 10.1016/j.lindif.2017.05.003
- Fonteyne, L., Eelbode, A., Lanszweert, I., Roels, E., Schelfhout, S., Duyck, W., & De Fruyt, F. (2018). Career goal engagement following negative feedback : influence of expectancy-value and perceived feedback accuracy. *International Journal of Vocational and Educational Guidance, 18*(2), 165–180. doi: 10.1007/s10775-017-9353-2
- Fonteyne, L., Wille, B., Duyck, W., & De Fruyt, F. (2017). Exploring vocational and academic fields of study : development and validation of the Flemish SIMON Interest Inventory (SIMON-I). *International Journal of Vocational and Educational Guidance, 17*(2), 233–262. doi: 10.1007/s10775-016-9327-9
- Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences, 14*(1), 49-66. doi:10.1016/j.lindif.2003.08.002
- Healy, C. C., & Mourton, D. L. (1983). Derivatives of the Self-Directed Search: Potential clinical and evaluative uses. *Journal of Vocational Behavior, 23*, 318-328.

- Hodara, M., & Lewis, K. (2017). How well does high school grade point average predict college performance by student urbanicity and timing of college entry? (REL 2017–250). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest. Retrieved 27th of March 2019 from <http://ies.ed.gov/ncee/edlabs>.
- Holland, J. L. (1963). Explorations of a theory of vocational choice and achievement. II. A four-year prediction study. *Psychological Reports*, 12(2), 547–594.
doi:10.2466/pr0.1963.12.2.547
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, FL: Psychological Assessment Resources.
- 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.
- Kriegbaum, Becker, & Spinath (2018). The relative importance of intelligence and motivation as predictors of school achievement: A meta-analysis. *Educational Research Review*, 25, 120-148. doi: 10.1016/j.edurev.2018.10.001
- Kuder, G. F. (1977). *Activity interests and occupational choice*. Chicago, IL: Science Research Associates.
- 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., Butt, S. M., Bradburn, J., & Prasad, J. (2018). Interests as predictors of performance: An omitted and underappreciated variable. *Journal of Vocational Behavior*, 108, 178-189. doi: 10.1016/j.jvb.2018.08.003

- 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
- Nye, C.D., Prasad, J., Bradburn, J., Elizondo, F. (2018). Improving the operationalization of interest congruence using polynomial regression. *Journal of Vocational Behavior*, 104, 154-169. doi: 10.1016/j.jvb.2017.10.012
- OECD (2017a). How Much Do Tertiary Students Pay and What Public Subsidies Do They Receive? Retrieved, October 23rd 2017, from <http://www.oecd.org/edu/> doi:10.1787/888932461199
- OECD (2017b). The State of Higher Education 2015-2016. Retrieved, June 6th 2019 from <http://www.oecd.org/education/imhe/the-state-of-higher-education-201516.htm>
- Päßler, K., & Hell, B. (2012). Do interests and cognitive abilities help explain college major choice equally well for women and men? *Journal of Career Assessment*, 20(4), 479–496. doi: 10.1177/1069072712450009
- Perera, H. N., & McIlveen, P. (2018). Vocational interest profiles: Profile replicability and relations with the STEM major choice and the Big-Five. *Journal of Vocational Behavior*, 106, 84-100. doi: 10.1016/j.jvb.2017.11.012
- Poole P., Shulruf B., Rudland, J., & Wilkinson T. (2012). Comparison of UMAT scores and GPA in prediction of performance in medical school: a national study. *Medical Education*, 46(2), 163-171. doi: 10.1111/j.1365-2923.2011.04078.x
- Prediger, D. J. (1982). Dimensions underlying Holland's hexagon: Missing link between interests and occupations? *Journal of Vocational Behavior*, 21, 259-287.
- Rayman, J., & Atanasoff, L. (1999). Holland's theory and career intervention: The power of the hexagon. *Journal of Vocational Behavior*, 55(1), 114–126. doi:10.1006/jvbe.1999.1701

- Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, 35, 83-92. doi:10.1016/j.intell.2006.05.004
- Roth, B., Becker, N., Romeyke, S., Schäfer, 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
- Rounds, J., & Su, R. (2014). The Nature and power of Interests. *Current Directions in Psychological Science*, 23(2), 98–103. doi:10.1177/0963721414522812
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. doi: 10.1037/0003-066 x.55.1.68
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. New York: Guilford Publishing.
- Schneider, B. (1997). The people make the place. *Personnel Psychology*, 40(3), 437–453. doi: 10.1111/j.1744-6570.1987.tb00609.x.
- Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, 143(6), 565-600. doi: 10.1037/bul0000098
- Shmueli, G. (2010). *To Explain or to Predict?* *Statistical Science*, 25(3), 289-310. doi: 10.1214/10-STS330
- Smart, J.C., Feldman, K.A. , & Ethington, C. A. (2000). *Academic disciplines: Holland's theory and the study of college students and faculty*. Nashville, TN: Vanderbilt University Press.
- Su, R., & Rounds, J. (2015). All STEM fields are not created equal: People and things interests explain gender disparities across STEM fields. *Frontiers in Psychology*, 6:189. doi: 10.3389/fpsyg.2015.00189

- 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
- Swanson, J. L., Hansen, J. I. C. (1988). Stability of vocational interests over 4-year, 8-year and 12-year intervals. *Journal of Vocational Behavior*, 33(2), 185-202. doi: 10.1016/0001-8791(88)90055-3
- Tett, L., Cree, V.E. & Christie, H. (2017). From further to higher education: transition as an on-going process. *Higher Education*, 73(3), 389-406. doi: 10.1007/s10734-016-0101-1
- 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., 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
- Wang, X. (2013a). Modeling Entrance into STEM Fields of Study Among Students Beginning at Community Colleges and Four-Year Institutions. *Research in Higher Education*, 54(6), 664-692. doi: 10.1007/s11162-013-9291-x

- Wang, X. (2013b). Why Students Choose STEM Majors: Motivation, High School Learning, and Postsecondary Context of Support. *American Educational Research Journal*, 50(5), 1081-1121. doi: 10.3102/0002831213488622
- Whitney, D. R. (1969). Predicting from expressed vocational choice: A review. *The Personnel and Guidance Journal*, 48(4), 279–286. doi:10.1002/j.2164-4918.1969.tb03318.x
- Wille, B., De Fruyt, F., & Feys, M. (2010). Vocational interests and Big Five traits as predictors of job instability. *Journal of Vocational Behavior* 76(3), 547-558. doi: 10.1016/j.jvb.2010.01.007
- World Economic Forum (2016). *The Human Capital Report*, retrieved 2 October 2018, from <http://reports.weforum.org/human-capital-report-2016/>
- Xu, Y. J. (2008). Gender Disparity in STEM Disciplines: A Study of Faculty Attrition and Turnover Intentions. *Research in Higher Education*, 49(7), 607-624. doi:10.1007/s11162-008-9097-4
- Young, G., Tokar, D. M., & Subich, L. M. (1998). Congruence revisited: Do 11 indices differentially predict job satisfaction and is the relation moderated by person and situation variables? *Journal of Vocational Behavior*, 52(2), 208–223. doi:10.1006/jvbe.1997.1587

Chapter 2. From Interest Assessment to Study Orientation: an Empirical Advice Set Engine

Published online, 3rd of June 2019.

*Stijn Schelfhout, Bart Wille, Lot Fonteyne, Elisabeth Roels, Filip De Fruyt &
Wouter Duyck (2019).*

*From Interest Assessment to Study Orientation:
An Empirical Advice Set Engine.*

The Journal of Experimental Education, DOI: 10.1080/00220973.2019.1605327

2.1. Abstract

Each student faces the challenge of choosing a study program that matches his or her vocational interest. A good person-environment fit (PE fit) between student and study program influences study success and persistence, prerequisites to obtaining the desired degree. But which criterion should be used when presenting advice sets of study options to orient students toward study programs that match their vocational interests? And how long should such a list of study options be? Moving beyond existing, non-evidence-based approaches, present study sets out to develop an empirical advice set engine (EASE) to optimize the process of matching future students to fitting study options. Compared to existing, non-evidence-based alternatives, EASE shows a better balance between the number and PE fit of the options presented. EASE may be a promising way to rethink how student PE fit information can be used in student orientation and higher education research.

2.2. Introduction

A student's vocational interest plays an important role in contemporary higher education. For instance, literature indicates that a good person-environment interest fit (PE fit) between a student and a study program predicts academic achievement and persistence (Allen & Robbins, 2010; Nye, Su, Rounds, & Drasgow, 2012; Rounds & Su, 2014; Tracey & Robbins, 2006; Tracey, Allen, & Robbins, 2012). As both academic achievement and persistence are prerequisites for graduation, students face an important decision when they have to choose a higher education study program in pursuit of the desired degree. Especially in educational set-ups with low admission fees or high scholarships, combined with open access to (nearly) all study programs, the possibilities toward higher education are almost limitless (OECD, 2017). As such, assisting students in their study choice by presenting them with a manageable list of study programs — also called an advice set — can provide substantial support. To generate such an advice set for an individual student, two factors need to be balanced. How many study programs should the advice set contain? And how high should the fit quality of the advice set be? Finding a balance between length and fit quality of the advice set would ensure that (prospective) students receive a manageable list of suitable study programs to choose from, while the list of programs still allows for study environment exploration (Holland, 1997). To our knowledge, educational research has not addressed these problems directly. In fact, educational research on the use of vocational interest and PE fit information toward study orientation has been scarce altogether. As a consequence, students, student counselors, and orientation tools often rely on tradition or non-evidence-based rules of thumb when establishing the length of an advice set.

The goal of the present study consists of introducing and exploring an empirical advice set engine (EASE). This EASE will generate an advice set of appropriate length and fit quality for each student, based on empirical data of students and study programs. At the base

of the model, we will use fine-grained methods to model the transition from a very good PE fit to a very bad one for each student. By establishing a critical point or threshold in the balance between the length and fit quality of the advice set, EASE will generate an advice set for each student. As such, we will explore how well our EASE methodology can balance length and fit quality of advice sets. Indeed, the model used will be refitted for each student, providing us with a measure of internal validation. We will also compare EASE to the more classic approaches using congruence indices, providing criterion validity at the student level (for an overview, see Camp & Chartrand, 1992). This comparison will give us an indication of the extent to which we could improve the quality of study orientation if we were to implement our engine. Finally, as validation of EASE at the program level, we will check to what extent successful students would receive their own study program as part of the EASE-generated advice set of study programs, without inflating the number of choices in this advice set.

The Importance of Vocational Interest in Student Orientation toward Higher Education

A definition of vocational interest usually incorporates a number of key features that enable and determine higher education study orientation: direction (or prediction), contextualization (interests have an object), stability, and motivation (Rounds, 1995; Rounds & Su, 2014; Su, Rounds, & Armstrong, 2009). First and foremost, vocational interests robustly predict study choice (Whitney, 1969). Today, the comparison of students' interests to study program environments has become a key element in study orientation. Holland proposed a model of vocational interest that enables such comparisons by using the same typology to represent students and study programs (Astin & Holland, 1961; Holland, 1997). This RIASEC typology takes the form of a clockwise hexagonal pattern containing six interest types or dimensions: realistic, investigative, artistic, social, enterprising, and conventional (Lippa, 1998). After decades of model development and evolution, this base

concept still remains highly influential, not in the least in the field of (higher) education (Nauta, 2010).

Second, interests are contextualized and always have an object, such as an activity or an environment (Rounds & Su, 2014). This means that students are interested in activities such as solving equations or translating a conversation or in environments wherein these activities take place, such as study programs (e.g., *mathematics* or *applied linguistics*) or future occupations (e.g., *mathematics teacher* or *interpreter*). When constructing interest questionnaires for study orientation, this object refers to individual study programs and their respective educational activities. Items and scales probing students' interests in these activities eventually lead to a student-specific-personal-interest (P) profile. Since the inception of the RIASEC model, literature has always harbored a vast set of instruments to determine such a P-profile (ACT, Inc, 2017; Arbona, 2000; Rayman & Atanasoff, 1999; SDS, 2017). For study orientation, such an instrument typically consists of a relatively large number of items covering the spectrum of human study-related behavior. SIMON-I, which was specifically designed for student transition toward a higher education setting, is a recent and validated example of this rich assessment tradition (Fonteyne, Wille, Duyck, & De Fruyt, 2017). Items of this instrument comprise both occupation titles (e.g., *linguist*, scored on the artistic scale) and (study-related) activities (e.g., *collecting quantitative and qualitative data* scored on the investigative scale) that one enjoys, to be scored on a dichotomous yes-no scale. The score on these items results in a personal RIASEC profile for each (future) student with (standardized) scores on all six dimensions, ranging from 0 to 100. A set of standardized RIASEC scores from SIMON-I will serve as the baseline from which to develop the EASE methodology in the present study. Apart from our specific study, EASE may just as well be applied to person profiles assessed by any Holland instrument other than SIMON-I.

However, before one can compare a student to a study program, the study program

profile has to be described using the same typology as the student's P-profile. Different approaches exist to describe an environment in terms of the RIASEC dimensions using, in contrast to a P-profile, an environment or E-profile. An often used approach in higher education research relies on the incumbent method (Holland, 1997). This method uses the assumption that a specific environment is determined through the people in the environment (Schneider, 1987). In other words, applying this assumption to a higher education setting, a study program is represented through its students. As such, the interest profiles of students occupying a certain study program environment (the so-called incumbents) are used to determine the interest profile of that study program environment. As an example from contemporary educational research, Allen and Robbins (2010) defined study programs in terms of the RIASEC dimensions by averaging out the RIASEC scores of students who demonstrated sufficiently high levels of academic achievement and persistence. By tracking a cohort of college freshmen throughout their study curriculum, Allen and Robbins (2010) showed that students with higher levels of congruence between their personal interests and the study program profiles (as determined through the incumbent method, based on historical data of their predecessors) had a better chance at obtaining their degree in a timely fashion.

This last example illustrates the importance of a third key feature of why vocational interest is so important toward higher education study orientation. Vocational interests are regarded as stable constructs (Low, Yoon, Roberts, & Rounds, 2005; Swanson & Hansen, 1988). Students who have a good match with their study program at the beginning of their higher education are likely to still have a good match when they graduate. This stability enables the possibility of researching the predictive power of vocational interest on study results of new students, based on their vocational interest and on historical results from graduates within a specific study program. For instance, recent meta-analytic research on almost 6,000 academic samples has indicated that vocational interests are moderately

correlated to variables indicating performance and persistence (Nye et al., 2012). Results also showed that especially the congruence between a person's vocational interest and his or her environment was of particular importance toward performance and persistence. This meta-analysis corrected historical views that doubted the influence of interests on performance variables because they focused largely on the absolute level of interest-dimension scores rather than on PE fit or congruence (Barrick & Mount, 2005).

The stability feature also enables the validation of study orientation. The attraction-selection-attrition model predicts that over time students will gravitate toward study programs that match their vocational interest (Schneider, 1987). This means that successful and persistent students become excellent incumbents for their (completed) study programs. As such, researchers can analyze existing or new methods of study orientation by investigating to which extent successful and persistent students would be oriented toward their original study choice made years ago. Such criterion validity is usually measured through a hit rate, with literature reporting numbers between 32% and 69%, depending on the interest inventory used (Burns, 2014; Donnay, 1997). Each match between a (successful and persistent) student's study program and the advice given through the method of study orientation is considered a hit for that study program. Derived from this hit rate, one could also investigate how many times a program was advised as part of an advice set. This alternative rate (or alt rate) for a study program will directly influence the length of an advice set. Indeed, if study programs have higher alt rates, students will receive advice sets with more study programs. However, one has to be careful not to inflate the future student's advice set with too many study programs in order to boost the validity and usability of the instrument. Such an expansion of the advice set could overwhelm the student with options and thus hinder the process of environment exploration. When validating an instrument for study orientation, one should therefore aim at high hit rates for all study programs, while keeping the alt rate for study

programs as low as possible. As an example, if the study program Economics has a hit rate of 81% with a 25% alt rate, it means that 81% of the students in this study program (Economics) would receive their own study program as a part of their advice set. This also means that 25% of the students inside and outside of Economics would receive this choice as a part of their advice set. In this study, we will explore to which extent the alt rate (in addition to the hit rate) provides extra information toward the validation of study orientation. Since both concepts are measured at study program level, the external validation of our EASE methodology will also be conducted at program level.

As a final characteristic, vocational interest can also act as motivation toward goal attainment, as described in social-cognitive theories of vocational interest (Lent, Brown, & Hackett, 1994; Rounds & Su, 2014). Indeed, interest in certain activities like solving equations or translating texts can (re)direct and energize the student's endeavors toward studying mathematics or applied linguistics, thus creating a study environment that facilitates focus on obtaining the desired degree. As such, the motivational component can explain why a good fit or match between student and study program leads to academic achievement and persistence.

Fitting Students to Study Programs

Early approaches to determine the PE fit between person (a student) and environment (a study program) profiles have long relied on the comparison of the highest scoring dimensions to obtain a congruence index, also called high point coding (Brown & Gore, 1994; Young, Tokar, & Subich, 1998). In such an index, the letters of the RIASEC dimensions for both P- and E-profiles are ranked from high to low based on the dimension scores. This procedure results in codes that describe students and programs such as SAIRCE or CESAIR. By comparing the rank and placement of the letters in the P- and E-profiles, most often only the first letters (one, two, or three dimensions at best), a categorical or ordinal measure of fit

is established. As an example, the Holland congruence index compares the highest dimensions of both P- and E-profiles (Holland, 1963). If these dimensions are the same (e.g., RIASEC versus RSIACE), the match between student and study program is deemed a good fit. Although these classic congruence indices have the advantage of being user friendly and transparent, they also have limitations (Tracey & Robbins, 2006). To give one example, too much emphasis is put on the absolute level of the scores, whereas the relative magnitude of the interest dimensions remains underused. For instance, both P- (60, 59, 59, 20, 30, 30) and E- (60, 31, 31, 20, 30, 30) profiles would result in an equivalent 3-letter code (RIA) based on classic congruence indexing. However, closer inspection of both profiles reveals substantial differences. The P-profile displays the highest score in the R dimension, with I and A being close seconds. In contrast, the E-profile displays a high R score, with the I, A, E, and C dimensions being at a much lower level. The previous example illustrates another problem. Letter coding does not provide a solution to tied dimensions (De Fruyt, 2002). Indeed, following the example above, the P- and E-profiles could also have been coded RAI, instead of RIA.

As a reaction to these concerns, alternative measures of PE fit have surfaced. One of these methods adopts a continuous approach, expressing the fit between person and environment through a mere correlation between profiles, while still being predictive of study success in the first year of higher education (Tracey et al., 2012). For instance, the PE correlation fit between a profile P (60, 59, 59, 20, 30, 30) and a profile E (60, 31, 31, 20, 30, 30) would amount to $r = .62$. The example clearly shows the difference with the letter-coding approach that coded both profiles as RIA without distinction. Indeed, by using a correlation, the relative magnitude of the dimension scores in both profiles is taken into account. The difference in elevation of the I and A dimensions of the profiles is reflected in a still high but

less than perfect correlation coefficient. This approach has the advantage that it uses the entire profile, while also rendering a continuous measure for additional, more fine-grained analyses.

The correlation approach is also immune to the absolute height of RIASEC dimension scores. Studies have shown that the average elevation of all dimensions does not have a direct effect on whether or not people want to engage in a certain occupation or activity (Prediger, 1998). However, literature also indicates that within lower elevated profiles the link between PE fit and results is even stronger (Darcy & Tracey, 2003; Tracey & Robbins, 2006). As such, study orientation should not focus on the height of the dimensions but on PE fit between profiles to avoid disadvantaging students with a low profile elevation. To address this problem, PE correlation fit seems a good solution. Finally, in addition to these advantages, the correlation index of PE fit is still easy to compute and interpret (from -1, being the worst fit possible, to +1, being a perfect fit) without much prior intensive data processing.

Translating PE fit information into study orientation advice

Despite the obvious theoretical and empirical advantages, questions remain regarding the practical implementation of this correlation approach to PE fit. Specifically, in a concrete study orientation situation, this approach generates a series of correlations between a student's interest profile and a set of available study options, reflecting the transition from a very good PE fit to a very bad one. Until now, we have no answer to the question concerning how good the PE fit between a student and a study program has to be before the program should be recommended to that specific student. This lack of a theoretically or empirically based objective criterion delineates a problem that the more classic congruence indices (see above) also could not solve. Indeed, educational literature has remained indecisive and vague concerning how the translation from PE fit to study orientation should be conceived. First, literature displays a multitude of congruence indices, all proposing different rules to indicate (the degree of) PE fit, each with its own (dis)advantages. As a result, what is deemed a good

fit is only valid within the confines of one specific index (Brown & Gore, 1994; Camp & Chartrand, 1992; Young et al., 1998); for instance, the dichotomous Holland index defines a good fit as a match between the highest dimensions (Holland, 1963). Second, none of these indices provides an answer to the question concerning how good exactly the PE fit between a student and a study program has to be before the program can be advised to that specific student. In other words, there is no objective and uniform criterion based on theory or empirical data that allows for making a distinction between a sufficient fit and an insufficient one. For instance, in the dichotomous Holland index described above, is it sufficient that only the highest dimensions match in order to include it in the advice set? Or do the second and third highest dimensions also need to match between student and program? As a result, contemporary study orientation still has to rely on mere tradition or suboptimal, non-evidence-based rules of thumb to guide students toward fitting study programs.

Present Study

How high does the fit between a student's interests and an available study program have to be to take this program into consideration as a potential study option, especially when comparing this fit to that of other study programs? To our knowledge, this question has not been researched in educational literature. Since there is no evidence-based criterion, the ideal length of a possible advice set featuring sufficiently high-fitting study programs also remains unknown. The objective of the present study is to answer both issues by balancing them against each other. As such, we will introduce and explore EASE, an Empirical Advice Set Engine. At its core, EASE optimizes the process of translating correlation PE fit information into concrete study advice, using an empirically fueled engine as a base for student friendly applications. Such a translation should always result in a balanced list of suggested study programs toward environment exploration, while containing only study programs that match a specific student's interests to a sufficient degree. To enable this translation, we will use a fine-

grained continuous method of PE correlation fit between a specific student and a list of available study programs, effectively modeling the transition from programs with a very bad PE fit, to programs with a very good PE fit. By building on this transition modeling, EASE will dispense a custom-made advice set of study options to each future student individually, while taking into account the correlation fit between the student's profile and the entire pool of available study options. As such, the criterion for this advice set will take the form of a minimal fit quality or threshold, relative to the available options. Study options that demonstrate a level of fit surpassing this threshold are included in the advice set, while the remainder of the study options is excluded, so that they do not have to be explored or processed by the student.

It is important to note that the decision to (not) include any given program in the advice set is always made relative to the pool of other possible study programs. As study orientation eventually leads to making a choice of study program, it is only fair that all possible choices be compared against the other choices. Ultimately, the proposed procedure should serve as the baseline for data-driven applications, while strengthening the quality and validity in establishing appropriate advice sets of study options for prospective higher education students. To this extent, the present study will explore three main research questions.

For the first question, we will test how well the novel method succeeds in balancing advice set length and fit quality for each student by using two large data sets containing student-interest measures. The first data set provides us with a large sample of real-life, successful students from various study fields, used to estimate study program interest profiles. The second data set provides us with a sample of future students seeking actual orientation toward fitting study options. As such we will test the following hypothesis,

H₁: EASE manages to balance the length and fit quality of student advice sets.

Since the balance between student's advice-set length and PE fit quality is a key feature of this study, we will also compare EASE to advice sets generated by classic congruence indices, such as the Holland index, discussed above, providing criterion validity at the student level.

H₂: EASE displays a better balance between student advice-set length and fit quality than classic congruence indices.

Finally, we will test the validity of the EASE methodology at the program level by exploring how many persistent and successful students would receive their own study program as part of their advice set. We also deem it worth investigating whether receiving the correct study program as part of the advice set does not needlessly inflate the length of the advice set. We will thus compare the EASE-generated advice sets to those generated through classic indices.

H₃: EASE generated advice sets have higher validity than those generated by classic congruence indices by displaying a better balance between hit rate and alt rate.

2.3. Method and Materials

Data Sets

All data were primarily gathered in function of a large, university-wide longitudinal project to enhance study orientation and study success among (future) students at a western-European university (Shanghai Top 100) with 11 diverse faculties. From this project, two obtained data sets were used, D₁ ($N_1 = 4,892$; 66% female) and D₂ ($N_2 = 7,063$; 61% female). D₁ features the scores on the RIASEC questionnaire SIMON-I (Fonteyne et al., 2017) from third bachelor's and master's assessed in the period between August 2013 and September 2015. Students were recruited from 62 study programs with on average 78 students for each program and a wide variety in student numbers ($SD = 80.20$). These students all met the conditions of academic success and perseverance by completing the first two years of their

study program (see Allen & Robbins, 2010). Only students who indicated that they would consider choosing the same study program again were included (97%). For each study program, the scores of all successful students or *incumbents* were averaged out, following the procedure of Allen and Robbins (2010). This operation provides us with an E-profile for each of the 62 study programs. These programs and their E-profile will function as possible study options for the current investigation. D₂ contains the RIASEC interest scores of future students (16 to 18 years old) on the verge of making the transition towards higher education. Interest assessments were conducted using a freely available, online version of SIMON-I in the period between January 2014 and September 2015 (see APPENDIX A). Highly irregular (for instance, scores of 0 and 100 on all dimensions) and incomplete profiles were excluded from the analyses (2%). All entries were rescaled analogous to D₁. There was no overlap between D₁ and D₂.

Procedures

EASE

Using the P-profiles of 7,063 future students and the E-profiles of 62 study programs, we will apply the EASE methodology to each student individually. As we are looking for a way to model the transition from a very good PE fit to a very bad one for each student individually, we have to correlate the student RIASEC profile (six dimensions) with each of the 62 study program RIASEC profiles (the same six dimensions). Such a correlation is a measure of PE *fit quality*. Table 1 shows an example for a single random student, ranking the fit quality of the student with the available study options from high to low. Each study option with a specific fit quality for an individual student is tied to a number of possible study options. This *options* variable indicates how big the advice set of the student would be, if the corresponding fit quality would act as the threshold (including all programs at or above its fit quality value) for making the advice set. Exploring the relation between the (PE) fit quality

and the number of options for this student even further, we observe a linear trend between both variables. This trend indicates that the distribution of fit quality within one student could approximate a uniform distribution, resulting in a gradual transition from very good to very bad fitting study options. We will test this approximation towards a uniform distribution for each student. Moreover, Figure 1 also shows that a high fit quality is tied to a low number of options and vice versa. This fit quality/options combination reflects the *balance* between the length and the minimal fit quality of the possible advice set for each student, formally defined as

$$balance = fit\ quality \times options \tag{1}$$

Balance has a single purpose: by finding the best possible balance for a student, we will be able to determine the optimal threshold for that student, weighing the number of study options against the minimal fit quality for study options in the advice set. As such, study options with a PE fit equal to or above this threshold will make up the proposed advice set.

Table 1. Balance between PE fit Quality and Number of Possible Study Options (length) of the Advice Set.

fit quality	options	balance
-0.87	62	-54.1
-0.85	61	-52.09
-0.83	60	-49.61
...
0.22	28	6.1
0.26	27	6.95
0.26	26	6.8
0.32	25	7.92
0.37	24	8.9
0.39	23	8.87
0.49	22	10.74
0.5	21	10.58
0.51	20	10.18
0.54	19	10.21
0.58	18	10.49

0.59	17	9.97
0.64	16	10.26
0.68	15	10.16
0.7	14	9.79
0.72	13	9.39
0.77	12	9.26
0.79	11	8.71
0.82	10	8.16
0.84	9	7.58
0.9	8	7.17
0,90	7	6.31
0.91	6	5.46
0.92	5	4.62
0.96	4	3.86
0.98	3	2.94
0.98	2	1.96
0.99	1	0.99

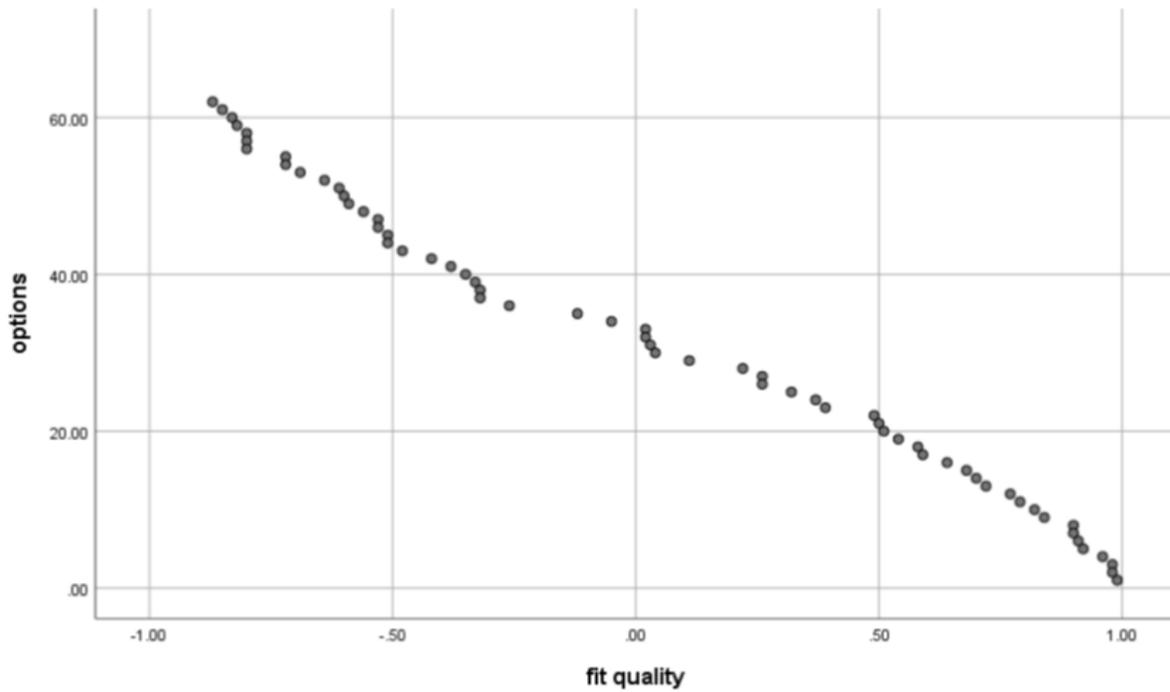


Figure 1. Scatterplot of Options and Fit Quality for a Random Student.

In order to find this optimal threshold, we introduce EASE. Its purpose will consist of finding the optimal threshold for each student separately, based on the balance variable.

Further inspection of Table 1 shows that the balance variable rises to more than ten and then

goes down again. In other words, it displays the larger part of a symmetrical and inverted, U-shaped curve, with the turnover point (the point where the rise stops and the descent begins) somewhere near the ten point mark of the balance variable. This turnover point is the equivalent of the *vertex* of a parabola and corresponds to our intended threshold. In other words, the vertex represents the point of ideal balance between a sufficient PE fit and an acceptable length of the advice set. If we can connect the vertex to the corresponding fit quality value, we have our threshold value for the advice set makeup. As we are looking for a PE correlation fit quality, and the balance variable displays an inverted U-shape curve, we propose a quadratic linear regression of fit quality on balance using the functional form of a parabola,

$$balance \sim a \times fit\ quality^2 + b \times fit\ quality + c + e \quad (2)$$

to model the balance curve. As such, parameters a , b and c need to be estimated while e represents the residual variance. Expressing fit quality as a function of balance allows us to estimate the x -coordinate of the vertex through its parameters by using

$$vertex = -\frac{b}{2a} = optimal\ correlation\ threshold \quad (3)$$

and as such obtain our optimal correlation threshold to reflect the ideal balance between length and fit quality of the advice set. All options above the computed optimal threshold are deemed of good enough fit quality and they will be included in the optimal and student-specific advice set. However, as the parabola is estimated through a regression, there will always be a margin of error. This margin of error could result in inflated or deflated

thresholds, illegitimately discarding or including study options to form the advice set. Considering we are advocating the principle of self-directed search, we deem it more important to keep borderline valid options in the advice set in opposition to discarding the less valid ones. To ensure EASE does not discard these valid options, we establish the actual threshold at the lower end of the threshold's confidence interval. Because the optimal threshold is based on parameter estimations, we use parameter confidence intervals (CI) to establish its own CI. In doing so, we take a conservative approach and use the upper and lower parameter bounds rendering the widest interval. Finally, the explained variance (R^2) of the quadratic regression provides us with a measure of how well the model fits the data. In other words, the EASE model fit will provide us with an estimate of how well the EASE methodology managed to balance advice set length and fit quality for a specific student. In sum, we define EASE as a quadratic linear regression, fueled by the model of a very good PE fit to a very bad PE fit between a student and a set of possible study programs, enabling the construction of an actual threshold for each individual student, which ultimately results in a balanced advice set of appropriate length and sufficient PE fit for each (future) student. In order for EASE to work, we do make the assumption that the PE fit values between a student and a pool of study programs entirely cover the correlation continuum. This assumption has to be tested.

Congruence indices comparisons.

As a test of the last two hypotheses, the EASE generated advice sets of study programs for each student are compared against more classic methods of constructing advice sets based on congruence indices, such as the letter congruency index discussed above. As these congruence indices all have specific features, we choose to include three classic indices, adapted or combined from the dichotomous first letter agreement index (Holland, 1963) and the two-letter agreement index (Healy & Mourton, 1983).

For the first comparison (H_2), the EASE data and letter method data are acquired from D_1 and D_2 . For the EASE data, the procedure is identical to the one described above. For the congruence indices using letter methods, i.e. $1L$, (one-letter), $2L$ (two-letter), and $1+2L$ (one- and two-letter combination) the procedure for making advice sets is conducted as follows. Study programs are included in the $1L$ advice set if the future student and study option profiles have the same highest scoring RIASEC dimension. For instance, a study program with E-profile code ECISAR (e.g., economics) would be included in the advice set of a student with P-profile ERCIAS. Study programs are included in the $2L$ advice set if the two highest dimensions from the study program profile reoccur in the three highest dimensions from the future student profile. For instance, a study program with E-profile code ECISAR (like economics) would be included in the advice set of a student with P-profile ERCIAS. Study programs are included in the $1+2L$ advice set if the conditions of $1L$ or $2L$ are met.

For the second comparison (H_3), the data is acquired from D_1 (successful and persistent students) and the procedure for both the EASE application and the letter methods is identical to the procedure from the first comparison, with one exception: the profiles of the (successful and persistent) students are also drawn from D_1 .

Measures and analyses

PE fit distribution

Before we apply EASE to the data, we have to verify to which extent the PE fit distribution for each student and the pool of study options approximates a uniform distribution. A good approximation indicates a gradual coverage of the correlation continuum, with a linear transition from very good to very bad fitting study options. The approximation is measured through an R^2 , as the result of a regression of options on fit quality (or vice versa).

EASE application

Figure 2 gives an example of an EASE application for a random student. As the

regression of the parabola model has to be carried out for each of the 7,063 students, analyses will report the range, mean and standard deviation across all students of the following variables: linear fit (R^2) (measuring how good the engine manages to balance the length and fit quality of the engine), the optimal and actual correlation threshold and the advice set size and average fit quality.

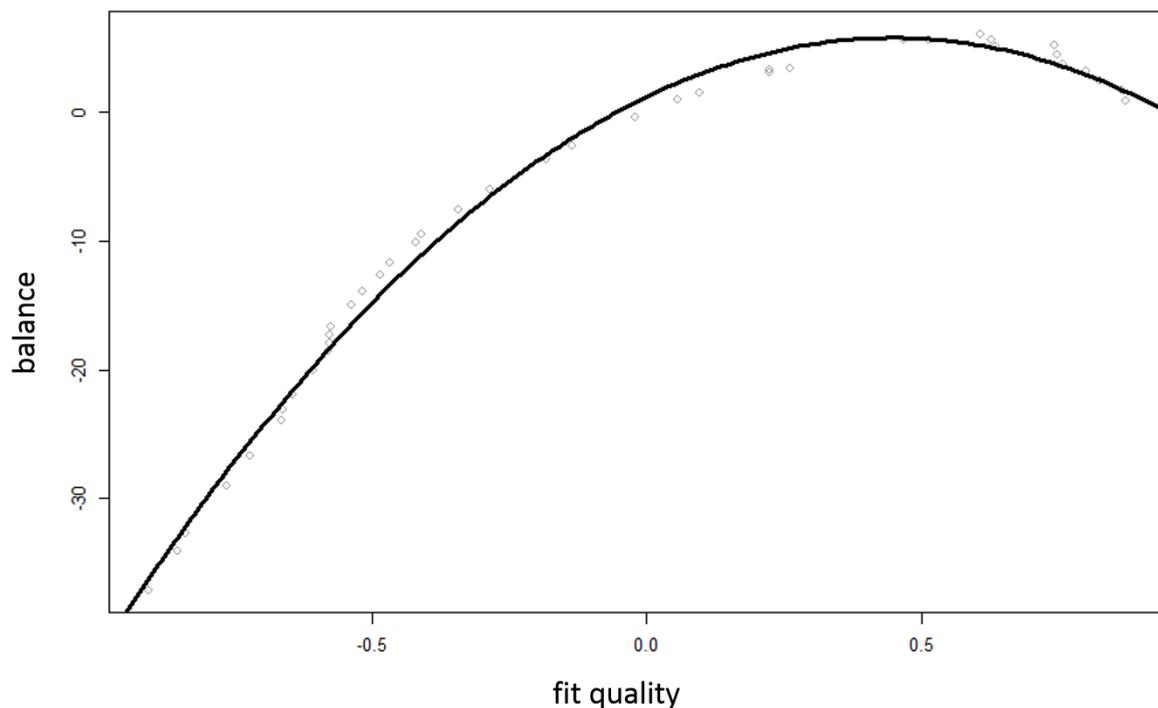


Figure 2. EASE Regression for a Random Student. Scatter plot data points are depicted in hollow. The quadratic regression is drawn in full.

EASE application versus classical 1L, 2L and 1+2L methods

Two comparisons are made. The first one will compare the balance between average advice set size and fit quality of both methods at the student level (H_2). The second comparison will compare the balance between the hit rate and alt rate of study programs (H_3). To control for the substantial differences in student numbers across study programs (see above), we use percentages (instead of absolute numbers) to ensure each study program has the same weight. For each comparison separately, the EASE results will be projected onto an interpolation of the results from the classic congruence indices (i.e., 1L, 2L and 1+2L

methods). As such, Figure 3 and Figure 4 show two polynomial interpolations, each consisting of two linear equations (depicted in full). These linear equations connect the results from 1L with 2L and 2L with 1+2L. Figure 3 depicts the relation between advice set size and fit quality of a student advice set. A congruence index method with a lower advice set fit quality is tied to a higher advice set size: students receive more study programs in their advice set, that consequentially fit worse. Figure 4 shows the relation between hit rate and alt rate of study programs. A congruence index method with a higher hit rate is tied to a higher alt rate. In other words, by increasing the number of options each student receives (alt rate), the chance rises they will also receive their own program as a part of the advice set (hit rate). By projecting the EASE results onto the interpolation of the classic methods (dotted line), hypothetical values can be established. We can now use a two-sided, one sample t-test to test whether these hypothetical values significantly differ from the observed EASE values to investigate if EASE (vs. classic congruence indices) indeed manages to obtain a better balance between the size and fit quality of a student advice set, and the hit rate and alt rate of study programs respectively. For an interpretation of the average differences, we will also report a Cohen's *d* effect size, with 0.01 = very small effect, 0.20 = small effect, 0.50 = medium effect, 0.80 = large effect, 1.20 = very large effect, 2.00 = huge effect (Sawilowsky, 2009).

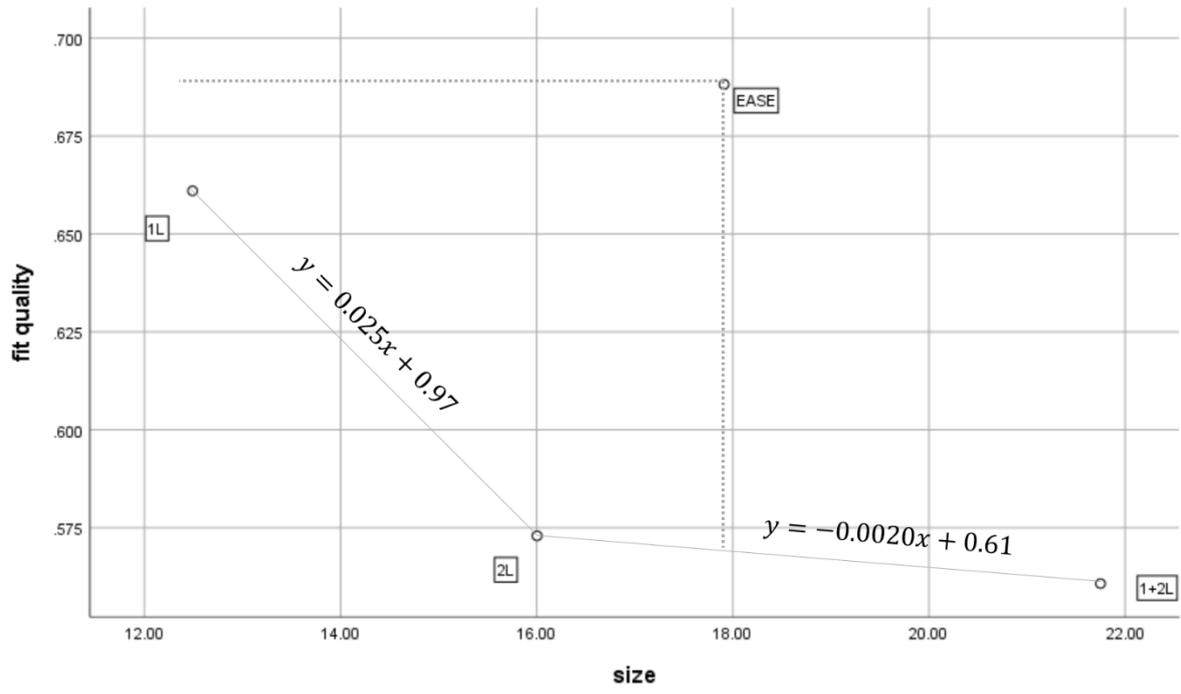


Figure 3. Projection of EASE on the Letter Method Interpolation of the Relation between the Size and Fit Quality of Student Advice Sets.

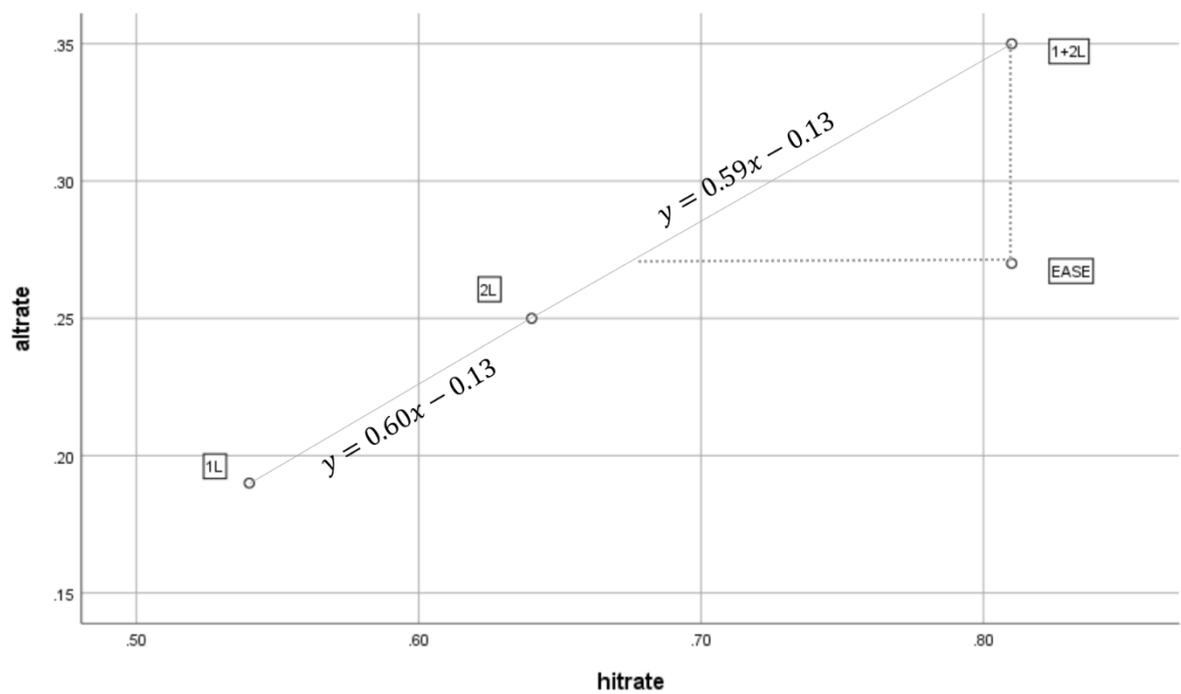


Figure 4. Projection of EASE on the Letter Method Interpolation of the Relation between Hit Rate and Alt Rate.

2.4. Results

PE fit Distribution

Figure 1 already hinted that the transition of PE fit within a student from a very good fitting study program to a very bad one is a very gradual and continuous process, following a uniform distribution. Formally, we tested this assumption for each prospective student, with an average regression R^2 across students of .97 ($SD = .03$), and a range from .84 to .99.

Hypothesis 1

Our first aim was to test how well EASE would be able to balance advice set length and fit quality. Our EASE methodology proved to be well capable of balancing advice set length and fit quality, as indicated by high levels of explained variance. Indeed, the quadratic regression of balance on fit quality resulted in a linear fit with an average R^2 of .99 ($SD = .01$), ranging from .86 to .99 across the (prospective) student sample. This high level of explained variance resulted in an accurate estimation of the optimal correlation threshold, which ranged from $r = .14$ to $r = .58$ ($M = .46$, $SD = .06$). The subsequent actual threshold ranged from $r = .11$ to $r = .53$ ($M = .44$, $SD = .06$). The width of its confidence interval ranged from .02 to .15 ($M = .06$, $SD = .02$). The advice sets were constructed based on the actual threshold for each prospective student separately. Figure 5 shows the distribution of the student advice set sizes. A student thus received an average advice set of almost 18 study options ($M = 17.91$, $SD = 5.37$), which is about 29% of the complete pool of 62 study options. Also, about 98% of the students received an advice set ranging from 7 to 28 options. Figure 6 shows the distribution of the (average) fit quality of the student advice sets. The fit quality of study options in an advice set was on average very high, with $M = .69$ ($SD = .09$) and a range in a right-skewed distribution from $r = .18$ to $r = .87$. Also, about 96% of the students had an advice set with an average fit quality of $r = .50$ or better. There were no advice sets with zero options. This means that all students received at least one possible study option as part of their advice set.

Considering the combined results of the analyses regarding our first hypothesis, we decide to accept H_1 .

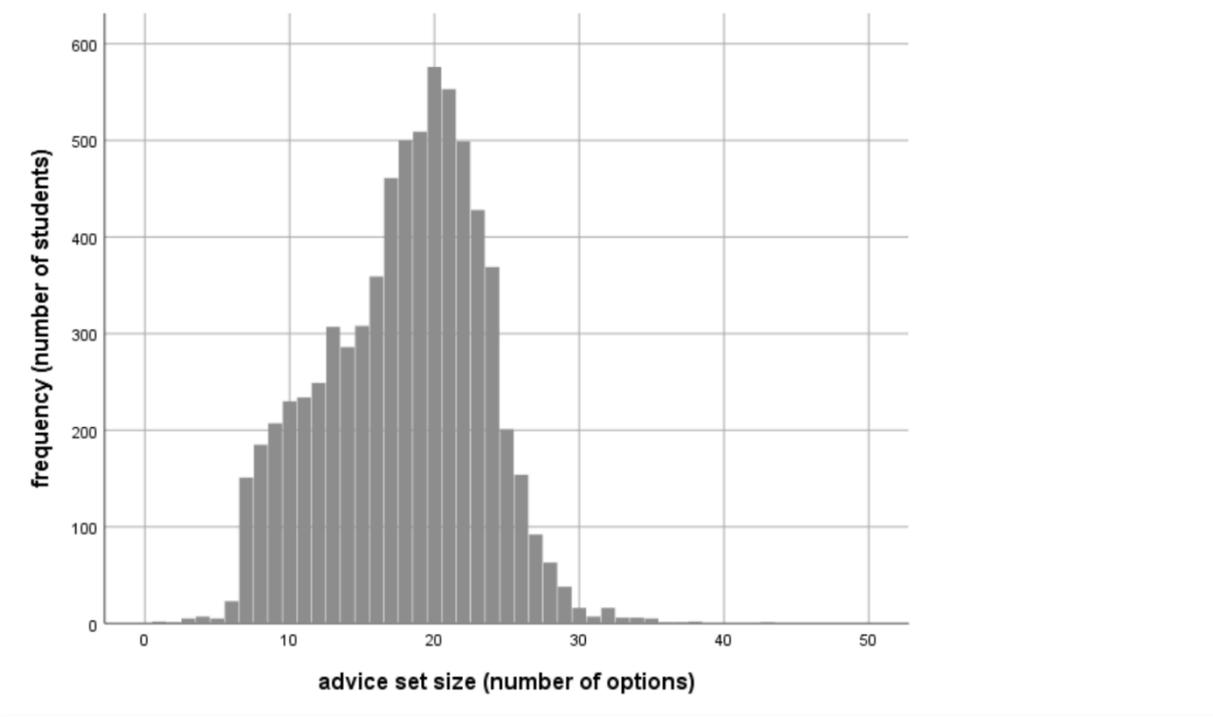


Figure 5. Distribution of Advice Set Size.

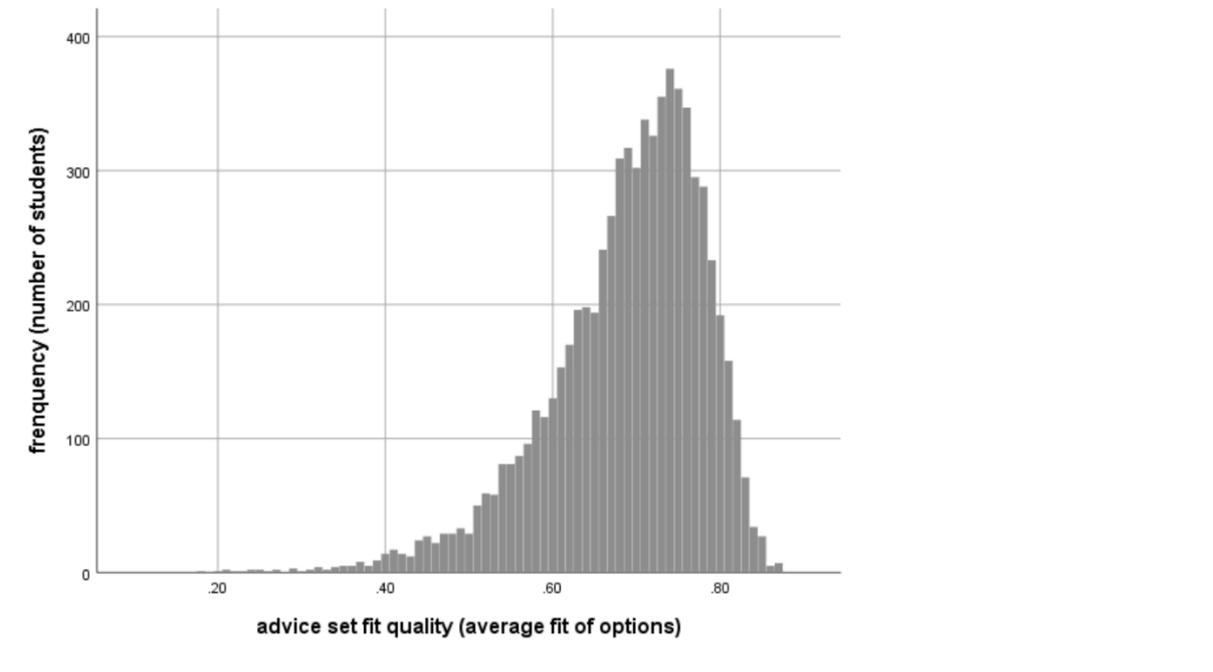


Figure 6. Distribution of Advice Set Fit Quality (based on the correlations between the P and E RIASEC profiles).

Comparison EASE and Congruence Indices

Hypothesis 2

Our second aim was to establish whether EASE displays a better balance between advice set size and fit quality than the classical approaches. Figure 3 clearly indicates the EASE results are above the interpolation line of the classic congruence indices. This deviation from the interpolation line already suggests that EASE manages to balance student study advice set size and program fit quality better than the classic congruence indices. By projecting the EASE values onto the interpolation line we can obtain hypothetical values to formally test the difference between EASE and the interpolation of the classic congruence indices on both advice set size and fit quality. Note that the congruence indices dispensed a varying number of zero sized advice sets (i.e., students who would receive no valid options). The 1L, 2L and 1+2L methods respectively rendered 401 (6%), 22 (< 1%) and 3 (< 1%) of such zero sized advice sets¹. Average fit quality values were computed by excluding the results of zero sized advice sets.

EASE generated student advice sets display an average size of 17.91 options. Inserting this value into the interpolation projects an average student advice set fit quality of $r = .57$. This is the fit quality that the classical letter methods would generate for an advice set size of 17.91. However, EASE generated student advice sets with a much better average fit quality of $r = .69$, compared to the interpolated $r = .57$. We can test this difference by using a two sided, one sample t-test. The difference between the observed EASE value and the projected interpolation value proved to be significant, $t(7062) = 106.54, p < .001$, Cohen's $d = 1.27$. This means that an equal advice set size for EASE and the classic congruence indices results in a very large difference in advice set fit quality, with EASE scoring $r = .12$ above the level of the interpolation line. Also, EASE generated advice sets display an average explained

¹ As an explanation, none of the 62 study programs had a RIASEC code starting with C, whereas students exist for which this is the dominant RIASEC dimension.

variance of 48% (i.e., $.69^2$), while the classic congruence indices only predict an explained variance of 32% ($.57^2$). In other words, EASE explains 16% more variance concerning the relation between a student's P-profile and his or her advice set of study programs (E-profiles) compared to the interpolation of the classic congruence indices (at equal advice set size).

By keeping advice set size constant, EASE yields better fit quality. It is also possible to reverse this rationale: does EASE generate larger advice sets, while still keeping a constant fit quality? EASE generated student advice sets display an average fit quality of $r = .69$. Inserting this into the interpolation equation yields a projection that is off the chart. As such, we propose to take a conservative approach and adopt the 1L edge value of 12.49 (advice set size) as an overestimation of the EASE projected value, as the actual interpolation would result in an even smaller sized advice set. However, EASE generated advice sets with a size of 17.91 options, compared to the interpolated 12.49. A two-sided, one sample t-test revealed a significant difference between this EASE advice set size and the conservative projection on the interpolation line, $t(7062) = 84.79, p < .001$, Cohen's $d = 1.01$. This means that EASE can maintain the same (maximal 1L) fit as the classic congruence indices, while rendering larger advice set sizes, with a difference of about 5.42 options. In sum, the tested projections of EASE on the interpolation line of the classic congruence methods indicate that EASE manages to outperform these classic methods in balancing advice set size and fit quality. As a consequence, we decide to accept H_2 .

Hypothesis 3

Our third aim was a test of the validity of the advice sets at the level of the study program. We investigated whether successful students received their own study program more often by using EASE over classic congruence indices through a higher hit rate, without inflating the advice set through a higher alt rate. Figure 4 clearly indicates the EASE results are below the interpolation line of the classic congruence indices. The deviation from the

interpolation line already suggests that EASE will have a higher hit rate at a lower alt rate. To formally test the difference between the classic congruence indices and the EASE methodology, we projected the observed EASE values onto the interpolation line of the classic congruence indices to obtain hypothetical values.

Through the EASE method, study programs receive an average hit rate of .81. Projecting this .81 on the interpolation line of the classic congruence indices renders an alt rate of .35. However, EASE displays an alt rate of .27. A two-sided, one-sample t-test revealed EASE has indeed a lower alt rate than the interpolation line, $t(61) = -6.04$, $p < .001$, Cohen's $d = 0.77$. This means that an equal hit rate of .81 for both EASE and the classic congruence indices, results in a large difference in alt rate, with EASE scoring .08 lower than the classic congruence indices. In other words, EASE improves (lowers) the alt rate of study programs with 23% compared to the classic congruence indices, at an equal hit rate of .81. In sum, through the use of EASE, programs have to appear less often in advice set to achieve the same hit rate.

Analogous to hypothesis 3, we can also reverse this rationale. What happens with the hit rate if we keep the alt rate constant? The EASE method generates an average alt rate in study programs of .27. Projecting this alt rate on the interpolation line renders a hit rate of .68. EASE however, displays a hit rate of .81. A two-sided, one-sample t-test revealed EASE has indeed a lower alt rate than the interpolation line, $t(61) = 8.82$, $p < .001$, Cohen's $d = 1.12$. This means that an equal alt rate of .27 for both EASE and the classic congruence indices results in a very large difference in hit rate, with EASE scoring .13 higher than the classic congruence indices. In other words, EASE improves (strengthens) the hit rate of study programs with 19% compared to the classic congruence indices, at an equal alt rate of .27. In other words, if one would present study programs equally often in study advice through both methods, EASE will yield a higher hit rate than classical methods. To summarize, the tested

projections of EASE on the interpolation line of the classic congruence methods indicate that EASE manages to outperform these classic methods in validity by demonstrating a better balance between hit rate and alt rate of study programs. As a consequence, we decide to accept H₃.

2.5. Discussion

Assisting (prospective) students in their study choice by orienting them towards a set of study programs that really matches their interest is of great importance to enhance study success and persistence in higher education (Tracey & Robbins, 2006; Allen & Robbins, 2010; Nye et al., 2012; Rounds & Su, 2014; Tracey et al., 2012). Until now, extant educational research remained indecisive and vague on how to translate PE fit into study advice. In the past, students, scholars and counselors relied on non-evidence based rules of thumb and a plethora of congruence indices, each with their own flaws and fortes, to establish goodness of fit (Brown & Gore, 1994; Camp & Chartrand, 1992; Healy & Mourton, 1983; Holland, 1963; Nye, et al., 2012; Young et al., 1998). Also, literature did not harbor an objective criterion to decide how well exactly the student's interests had to match a study program in order for the program to be eligible as a part of the advice set of study programs presented to a specific student. As a consequence, the ideal length of such a custom made advice set also remained unknown. This crux in educational literature is quite surprising as we have argued that vocational interest and PE fit are of capital importance towards higher education study orientation through the features of prediction, contextualization, stability and motivation (Lent et al., 1994; Low et al., 2005; Nauta, 2010; Rounds, 1995; Rounds & Su, 2014; Swanson & Hansen, 1988; Su et al. 2009; Whitney, 1969). In order to translate PE fit into study advice, the present study proposes the EASE (*Empirical Advise Set Engine*) methodology. EASE empirically generates an individualized advice set of study programs that is sufficiently large and of good fit quality for each future student. In doing so, the engine

balances the number of study programs in the advice set versus the minimal fit quality required for such a study program to enter the advice set. At its base, EASE uses the benefits of the fine-grained PE correlation fit measure to model the transition from a very good PE fit to a very bad PE fit between any given student and a set of study options (Allen & Robbins, 2010; Tracey et al., 2012). By finding the ideal balance between the number of study options and minimal PE fit, a correlation threshold is generated for each student. Study programs with a PE fit (regarding the specific student) above the threshold are added to the advice set and presented to the student as part of the final advice set, while the other options are no longer taken into account as programs fitting the student's interests.

To explore the possibilities of our EASE methodology, we presented three research questions. (1) How well does the EASE methodology succeed in balancing the length and fit quality of a student advice set? (2) How do the EASE generated advice sets compare to sets generated with more traditional congruence (letter) indices? (3) How valid is the EASE methodology?

As an answer to the first question, EASE displays a remarkable ability to balance length and fit quality of student advice sets by determining an empirical PE fit threshold for each individual student, through the use of the parabola model. This threshold leads to varied student advice sets of about 18 study programs, with about 98% of the prospective students receiving an advice set between 7 and 28 choices, leaving ample room for study environment exploration (Holland, 1997; Gottfredson & Holland, 1975). Our study also includes a number of validation mechanisms for the parabola model. Indeed, the model fits to all student profiles individually, while also providing two forms of criterion validity at the student and study program level, addressed in research questions two and three respectively.

Indeed, as an answer to the second question, EASE presents student advice sets that are qualitatively better than those generated with the classical congruence indices. For

instance, an EASE advice set of about 18 study programs delivers study advice to future students that explains 48% of the variance in the relation between the student's P-profile and the advice set of study programs. This variance level is 16% higher than the level achieved by the classic congruence indices. Also, about 98% of all students received an advice set with a fit quality of $r = .50$ or better.

And finally, as an answer to the third question, our EASE methodology shows strong criterion validity for study programs; about 81% of all successful students received their own study choice as part of their EASE generated advice set. Comparing the EASE hit rate to the range reported in literature (i.e., 32% to 69%), our method seems to be more accurate than using classic methods of making the PE fit (Burns, 2014; Donnay, 1997). Moreover, EASE also outperforms a combination of congruence indices by displaying a hit rate that is 19% higher at equal alt rates (Holland, 1963; Healy & Mourton, 1983). The results from this last question also show the incremental research value of the alt rate when validating study orientation tools. For sure, high hit rates in study orientation are important to ensure validity, but not at all cost. Good study orientation should also monitor whether the alt rates are not needlessly inflating the student's advice set: if too many less fitting programs are suggested, the process of environment exploration will suffer (Holland, 1997; Gottfredson & Holland, 1975). Classic congruence indices may present a strong hit rate or a low alt rate. But EASE has a better balance between both, with an alt rate that is 23% lower than those of the classic congruence indices, measured at equal hit rates. As such, EASE presents the right programs to the right students, without having to present programs too often to achieve that.

In sum, the exploration of our three research questions clearly shows the classical congruence indices still produce acceptable results concerning fit quality and validity of the generated advice sets in order to orient students towards higher education. However, when comparing these results to those obtained through the EASE methodology, EASE provides

each student individually with better fitting and valid study options. On top of better quality and validity of the advice set, the EASE methodology also presents a number of positive features that the congruence indices fail to reproduce. Indeed, when generating advice sets for future students, EASE ensures an orientation advice set of at least one study option for each student, while the congruence indices cannot provide orientation for up to 6% (i.e. the number of zero-sized advice sets for the 1L method) of all future students. As such, EASE has a better answer (vs. the classic methods) to the absence of programs with a dominant C-dimension. Next, EASE succeeds in establishing an objective, data driven and student specific criterion that allows to identify study programs that should be part of the student specific advice set orientation (vs. study programs that should be discarded). Finally, EASE establishes this criterion while comparing all available study programs against each other. This comparison seems only fair when considering study orientation should ultimately lead to making a choice between study programs.

Theoretical Implications

As an important theoretical addition to the structure of PE fit, we established that the transition from a very good PE fit to a very bad PE fit is apparently a very gradual and continuous process for each individual student. The correlation approach thus provides a continuous, fine-grained measure for modeling the structure of PE fit as an approximated uniform distribution. This also means that the parabola estimated for each individual student has properties that find their origin in the uniform distribution of PE fit. Though these properties were not intended as such, they are a direct consequence of the empirically observed PE fit distribution and they will influence the length and quality of the advice set. For instance, EASE uses the symmetry about the parabola vertex to make the distinction which programs are suitable for the student and which are not. Moreover, this distinction is made more clear cut as the programs are gradually distributed across the course of the

parabola. However, EASE does not use the full course of the parabola. Indeed, EASE does not aim for advice sets equal in length to everyone. Instead, each student should receive a list of options based on the fit of his profile to the pool of available programs. By using only a part of the parabola course including the vertex, EASE also succeeds in balancing the number of suggested options. As such, EASE renders advice sets of study options that are large enough for the intended self-exploration, without inflating the advice sets to unworkable lengths.

Moreover, this research does not have to limit itself to the domain of education. We speculate that the structural uniform distribution of PE fit as shown in this study could also be present in other settings like work or hobbies, effectively paving the way for the introduction of EASE in these settings as well. As such, we advocate further research on the distribution of PE fit between a student and study programs. We also advise to always explore this uniform distribution when using the EASE methodology.

Practical Implications and Limitations

The analyses above show that EASE offers a good method to offer prospective students a list of suggested programs, that is not too short or long, and that fits their interests well. The practical implications for study orientation towards higher education of the proposed methodology are tied to a number of boundary conditions that deserve further attention. A first and obvious requirement is that there are interest profiles of both (future) students and study programs. Although in the present study these two types of profiles resulted from the same interest instrument (i.e. SIMON-I, Fonteyne et al., 2017) which was administered to future students as well as successful students, this is not essential. The only prerequisite is that both personal and environmental profiles are commensurate measures, consisting of the same number of conceptually related (e.g., RIASEC) dimensions and thus it both mathematically possible and conceptually reasonable to compute the correlations

between both profiles as a commensurate assessment. It should be clear from the above that these correlations form the basis of the EASE method. Any assessment can make use of the EASE methodology, as long as the compared measures of person and environment are commensurate.

A second requirement consists of a sufficiently large pool of study options. The present exploration already showed that a set of 62 options is sufficient to extract a very stable advice set. The high amounts of explained variance in the EASE application do seem to suggest that even a smaller pool of study options could enable balancing advice set length and fit quality. The question remains how small the pool of study options can become while still keeping the PE correlation fit continuum sufficiently covered. This needs to be clarified in future research, while at the same time asking the question if such a small pool needs an advice set to begin with.

A third requirement is of course a (set of) student RIASEC profiles to apply EASE and generate advice sets. For individual student orientation, the data of a single future student is sufficient to construct a distribution pattern and apply our EASE to that specific student generating a valid and precise advice set, containing an appropriate and sufficient number of study programs.

A final requirement consists of (data fueling) EASE itself. In this exploration, we have provided but one possible configuration, defining balance as the (simple) product of options and fit quality in order to pinpoint a correlation threshold. Other setups might require adaptations like weighting the components, if one or the other would be more important in a specific context.

Future Research and Applications

EASE has the potential to fuel an orientation tool for centers of higher education like colleges or universities that harbor an extensive set of (diverse) study programs. Automating

this engine through an online application can reach a vast number of (future) students to fill out any RIASEC questionnaire. This will enable the entire EASE procedure by meeting the mentioned requirements. By featuring any RIASEC questionnaire, data can be gathered on the profiles of both actual and future students. Actual students will act as incumbents effectively rendering study program profiles, while the profile of future students is run through the engine to generate appropriate advice sets towards study choice. Advice sets can take a form similar to Table 1, listing appropriate programs instead of a number of options, (not) including the PE fit through a fit quality for purposes of further exploring the advice set. We also refer to APPENDIX B that contains a full example for one student, featuring both the EASE execution code and practical application of the algorithm.

Results from the current set of analyses already suggest that the presented EASE methodology has the potential to significantly advance our understanding of the concept of PE fit and how it can be applied in practice. As such, it would also be highly beneficial to use these data from automated online applications to facilitate this process of ongoing research. Indeed, additional research on this method is desired, especially on the properties of EASE across different instruments and contexts. A correlation fit can be used independently of the featured instrument, as long as it is possible to establish a correlation between a profile P and E. In theory, this makes our method appropriate for SIMON-I, UNIACT, Self - Directed Search or any other Holland-based instrument as long as it features all six RIASEC scales (ACT, 2017; Arbona, 2000; Fonteyne et al., 2017; Gottfredson & Holland, 1975; Nauta, 2010; Rayman & Atanasoff, 1999; SDS, 2017). It is worthwhile to compare said instruments on variables such as fit quality and advice set size.

Similarly, EASE offers the ability to explore to which extent and under which form the EASE method can be applied to contexts other than education as the results from the uniform distribution approximation seem to indicate. For instance, given the centrality of

interests in many aspects of professional career development, we deem it worthwhile to examine to which extent this threshold method may also be applied in actual working contexts. The EASE method could help in generating advice sets consisting of job profiles which can then be linked (e.g., by labor agencies) to the interest profiles of individual job seekers.

Conclusion

Person-environment interest fit is an important predictor of higher education performance and persistence. Nevertheless, little progress has been made over the past years in charting student PE fit distribution and in developing methodologies to translate PE fit information into valid and workable study advice. The method proposed in the current work introduces a novel way of translating PE fit into student orientation. Compared to more traditional and mainly convention-based congruence index approaches to PE fit and study orientation, this new methodology ensures the creation of advice sets, balanced in length (to enable environment exploration) and fit quality (in terms of correlation PE fit). In sum, EASE may be a promising way to rethink how student PE fit information can be used in both fundamental research and practical applications regarding student orientation and higher education research.

2.6. References

- ACT, Inc. (2017). The ACT Interest Inventory technical manual. Retrieved October, 23, 2017, from <http://www.act.org/content/dam/act/unsecured/documents/>
- Allen, J. & Robbins, S. (2010). Effects of Interest–Study program Congruence, Motivation, and Academic Performance on Timely Degree Attainment. *Journal of Counseling Psychology, 57*(1), 23-35. doi:10.1037/a0017267
- Arbona, C. (2000). Practice and research in career counseling and development – 1999. *Career Development Quarterly, 49*(2), 98-134. doi:10.1002/j.2161-0045.2000.tb00554.x
- Astin, A. W., & Holland, J. L. (1961). The environmental assessment technique: A way to measure college environments. *Journal of Educational Psychology, 52*(6), 308–316. doi: 10.1037/h0040137
- Barrick, M. R., & Mount, M. K. (2005). Yes, personality matters: Moving on to more important matters. *Human Performance, 18*(4), 359–372. doi: 10.1207/s15327043hup1804_3
- Brown, S. D., & Gore, P. A. Jr. (1994). An evaluation of interest congruence indices: Distribution characteristics and measurement properties. *Journal of Vocational Behavior, 45*, 310-327. doi:10.1006/jvbe.1994.1038
- Burns, S. T. (2014). Validity of Person Matching in Vocational Interest Inventories. *The Career Development Quarterly, 62*(2), 114-127. doi: 10.1002/j.2161-0045.2014.00074.x
- Camp, C. C., & Chartrand, J. M. (1992). A comparison and evaluation of interest congruence indices. *Journal of Vocational Behavior, 41*, 162-182. *Counseling Psychology, 45*, 204-211. doi:10.1037/0022-0167.45.2.204

- Darcy, M., & Tracey, T.J.G. (2003). Integrating Abilities and Interests in Career Choice: Maximal versus Typical Assessment, *Journal of Career Assessment*, *11*(2), 219-237. doi: 10.1177/1069072703011002007
- De Fruyt, F. (2002). A person-centered approach to P–E fit questions using a multiple trait model. *Journal of Vocational Behavior*, *60*(1), 73–90. doi: 10.1006/jvbe.2001.1816
- Donnay, D. (1997). E. K. Strong's legacy and beyond: 70 years of the Strong Interest Inventory. *The Career Development Quarterly*, *46*(1), 2–22. doi: 10.1002/j.2161-0045.1997.tb00688.x
- Fonteyne, L., Wille, B., Duyck, W., & De Fruyt, F. (2017). Exploring vocational and academic fields of study : development and validation of the Flemish SIMON Interest Inventory (SIMON-I). *International Journal of Vocational and Educational Guidance*, *17*(2), 233–262. doi: 10.1007/s10775-016-9327-9
- Gottfredson, G.D., & Holland, J.L. (1975). Vocational choices of men and women – comparison of predictors from self-directed search. *Journal of Counseling Psychology*, *22*(1), 28-34. doi:10.1037/h0076150
- Gottfredson, G.D., & Holland, J.L. (1991). *The Position Classification Inventory: Professional manual*, Odessa, FL: Psychological Assessment Resources.
- Healy, C. C., & Mourton, D. L. (1983). Derivatives of the Self-Directed Search: Potential clinical and evaluative uses. *Journal of Vocational Behavior*, *23*, 318-328.
- Holland, J. L. (1963). Explorations of a theory of vocational choice and achievement. II. A four-year prediction study. *Psychological Reports*, *12*, 547-594.
- Holland, J.L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, FL: Psychological Assessment Resources.
- Lent, R. W., Brown, S. D., Hackett, G. (1994). Towards a Unifying Social Cognitive Theory of Academic Interest, Choice and Performance. *Journal of Vocational Behavior*, *45*(1), 79-122. doi: 10.1006/jvbe.1994.1027

- Lippa, R. (1998). Gender-related individual differences and the structure of vocational interests: the importance of the people-things dimension. *Journal of Personality and Social Psychology, 74*(4), 996-1009. doi:10.1037/0022-3514.74.4.996
- Low, K. S. D., 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), 713–737.
- 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 (2017). Retrieved from <http://www.oecd.org/edu/> doi:10.1787/888932461199
- Prediger, D. J. (1998). Is interest profile level relevant to career counseling? *Journal of Counseling Psychology, 45*(2), 204-211. doi:10.1037/0022-0167.45.2.204
- Rayman, J., & Atanasoff, L. (1999). Holland's theory and career intervention: The power of the hexagon. *Journal of Vocational Behavior, 55*(1), 114-126.
doi:10.1006/jvbe.1999.1701
- Rounds, J. B. (1995). Vocational interests: Evaluation of structural hypotheses. In D. Lubinski & R. V. Dawis (Eds.), *Assessing individual differences in human behavior: New concepts, methods, and findings* (pp. 177–232). Palo Alto, CA: Consulting Psychologists Press.
- Rounds, J., & Su, R. (2014). The Nature and power of Interests. *Current Directions in Psychological Science, 23*(2), 98-103. doi: 10.1177/0963721414522812

- Sawilowsky, S (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, 8(2), 467-474.
- Schneider, B. (1987). The people make the place. *Personnel Psychology*, 40, 437-453.
doi:10.1111/j.1744-6570.1987.tb00609.x
- SDS (2017). Self-Directed Search, a part of Aspira 360. Retrieved October, 23, 2017, from
<http://www.self-directed-search.com/>
- Swanson, J. L., & Hansen, J. I. C. (1988). Stability of Vocational Interest over 4-year, 8-year and 12-year Intervals. *Journal of Vocational Behavior*, 33(2), 185-202.
- 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
- Tracey, T.J.G., & Robbins, S.B. (2006). The interest-study program congruence and college success relation: A longitudinal study. *Journal of Vocational Behavior*, 69(1), 64-89.
doi:10.1016/j.jvb.2005.11.003
- 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
- Whitney, D. R. (1969), Predicting from expressed vocational choice: a review. *The Personnel and Guidance Journal*, 48(4), 279-286. doi:10.1002/j.2164-4918.1969.tb03318.x
- Young, G., Tokar, D. M., & Subich, L. M. (1998). Congruence revisited: Do 11 indices differentially predict job satisfaction and is the relation moderated by person and situation variables? *Journal of Vocational Behavior*, 52, 208-223.
doi:10.1006/jvbe.1997.1587

2.7. APPENDIX A:

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.

Dimensions were masked for the participant.

	Dimension	YES	NO
Developing electronic systems	R		
Analysing the grammatical structure of a sentence	I		
Helping people with speech disorders	S		
Recruiting a job candidate	E		
Monitoring the quality standards for food safety and hygiene	C		
Repairing malfunctioning electrical equipment	R		
Carrying out laboratorial analyses	I		
Designing a poster for an exhibition	A		
Helping others with their personal problems	S		
Organising a conference	E		
Preparing financial reports	C		
Being responsible for the maintenance of IT hardware	R		
Analysing statistics	I		
Designing webpages	A		
Developing council prevention campaigns	S		
Presenting new policy propositions	E		
Collecting quantitative and qualitative data	I		
Develop new methods for industrial production	R		
Treating diseases in animals	I		
Editing the sound and images for a movie	A		
Formulating education and training policies	S		
Drawing up the budgets	C		
Doing the follow up on building sites	R		
Analysing x-rays/brain scans	I		
Fit out a show room	A		
Sport guidance for children, the elderly, ...	S		
Formulate a theory about the differences between population groups	I		
Monitor quality standards	C		
Maintaining airplanes	R		
Investigating the impact of historical people	A		
Composing a work of music	A		
Providing guidance for victims	S		
Selling a product or service	E		
Calculating prices	C		
Installing and maintaining computer servers	R		
Designing an advertising folder	A		
Providing information about the assistance for the poor	S		

Drawing up an organisational business or policy plan	E
Checking bank transactions	C
Developing windmill parks	R
Prove a theorem	I
Analysing text structures	A
Giving travel advice	S
Negotiating contracts	E
Drawing up a contract	C
Investigating chromosomal defects	I
Writing scenarios	A
Holding tests, questionnaires and in-depth interviews	S
Screening the administration	C
Working on a drilling rig	R
Turning an idea into a film	A
Giving care to patients	S
Restructuring an organisation or company	E
Checking the compliance of regulations	C
Excluding alternative explanations through experiments	I
Designing the layout of a hospital	A
Advising youngsters regarding their vocational choice	S
Exploring new economic markets	E
Drawing up the annual report	C
Setting up a festival stage	R
Developing a new medicine	I
Writing a review	A
Giving training in communication skills	S
Starting up an enterprise	E
Investigating a cost structure	C
Creating a technical drawing	R
Putting theories in their historical and social context	I
Creating an art piece	A
Giving health advice	S
Giving health and parenting education	E
Calculating expenses	C
Disassembling electrical appliances	R
Comparing cultures	A
Guiding minority groups on the job market	S
Conducting a meeting	E
Drawing up a timetable	C
Measuring a lane	R
Supporting and following up foster families	S
Attracting sponsors	E
Standing in front of a classroom	S
Leading a team	E
Managing a database	C
Collecting soil samples	R

Beginning a herbarium (a plant collection)	I
Counseling underprivileged people	S
Formulating a treatment plan	S
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.

Dimensions were masked for the participant.

	Dimension	YES	NO
Industrial designer	R		
Civil engineer	I		
Fashion designer	A		
Policy advisor in political and international relations	E		
Recruitment and selection advisor	E		
Damage expert	C		
Agricultural technician	R		
Teacher	S		
Business economist	C		
Accountant	C		
Electrical engineer	R		
Biologist	I		
Art/music teacher	A		
Speech therapist	S		
Bank manager	C		
Landscape architect	R		
Physicist	I		
Editor	A		
Student counselor	S		
Tax supervisor	C		
Neurologist	I		
Policy advisor art and culture	A		
Educator	S		
Marketing manager	E		
Safety advisor	C		
Construction manager	R		
Historian	I		
Director	A		
Communication manager	E		
Manager (of a company)	E		
Judge	C		
Forester	R		
Researcher	I		
Graphic designer	A		
Psychologist	S		

Lawyer	E
Notary	C
Mathematician	I
Art historian	A
Social worker	S
Politician	E
Pilot	R
Pharmacist	I
Linguist	A
Divorce mediator	S
Journalist	A
Structural engineer	R
Lab assistant	I
Photographer	A
Nurse	S
Advertising campaign manager	E
Chemist	I
Tax specialist	C
Architect	R
Artist	A
Educational scientist	S
Librarian	A
Philosopher	I
Representative	E
Geneticist	I
Interior designer	A
Estate agent	E
Physiotherapist	S
Meteorologist	I
Sales manager	E
Statistician	I

2.8. APPENDIX B:

EASE Manual and Executable RStudio Code: an Example for one Student

1. Obtain the RIASEC profile of the student. For instance, $R = 17$, $I = 59$, $A = 12$, $S = 14$, $E = 0$, $C = 9$. For this example, scores are scaled from 1 to 100. Any vocational interest instrument (or scaling) can be used, as long as it covers all six RIASEC dimensions.
2. Obtain the RIASEC profiles of the programs. For instance, a bachelor in mathematics has a profile of $R = 25.98$, $I = 41.01$, $A = 28.52$, $S = 24.81$, $E = 26.04$, $C = 33.03$. Scores are scaled from 1 to 100. Again, any instrument (or scaling/method) can be used, as long as it covers all six RIASEC dimensions. For this example, we used an incumbent method as described by Allen and Robbins (2010) to construct 62 program profiles.
3. Obtain the correlation between the student and each of the program profiles. For instance, correlating the profiles from the student and mathematics program yields $r = .83$. For this example, we correlated the student profile to each of the 62 program profiles. For obtaining this multitude of correlations, we advise using an excel worksheet. Insert student profile and mathematics profile in one horizontal row; you can add a horizontal identifier and column headings if so desired. For instance, cells A1 to F1 contain the student RIASEC profile, and cells H1 to M1 contain the mathematics RIASEC profile. Enter the code for the correlation in cell G1, “=CORREL(\$A1:\$F1;H1:M1)”.
Enter the code for the correlation in cell G1, “=CORREL(\$A1:\$F1;H1:M1)”.
4. Add the next program profile using the same procedure. As such, leave one space N1 for the correlation code and insert the RIASEC dimension scores of the program in cells O1 to T1. Copy and paste the correlation code from G1 to N1. The fixed student profile cells, indicated by a dollar sign, will remain in place, while the profile cells will change from H1:M1 to O1:T1. Repeat the procedure for all programs. For future reference, adding new student profiles can be done by using a new row for each student. Add student profiles using columns A to F. Copy the remainder of the program data (and the correlations) by

selecting the top row and double right clicking on the lower right corner of your selection.

To make a file only containing the correlation values, copy the values (not the formulas) of the entire file (sheet) to a new file and delete everything but the correlations (or optional identifiers).

5. For the present example, we have listed all programs and their correlations to the student profile at the end of this appendix. Create the excel file “onestudent” by pasting the 62 correlations (with optional programs as column titles) in cells A1 to BJ1 of an excel sheet. Name the excel file “onestudent”. The correlations do not need to be ranked.
6. Import the excel datafile “onestudent” into R (Studio). As you import the code, do indicate if your file contains column titles. Paste the EASE executable code (see below) into an R(studio) script, load the packages mentioned and follow instructions where needed. Run the EASE R-script with “onestudent” as input. We have annotated the code with editorial comments to clarify the application and to link this application to the EASE paper. Comments are annotated in bold and preceded by #.

EASE executable code

load packages broom, lsr, psych, import excel sheet onestudent

for other datasets, simply replace onestudent with the name of your datafile and load

the file

declaring matrices for matters of easier programming

mydata = onestudent

mydata = as.matrix(mydata)

Dprep = mydata

```
Dprep = as.matrix(Dprep)
```

```
Dprep = t(Dprep)
```

```
testmatrix = mydata
```

```
testmatrix = as.matrix(testmatrix)
```

```
# declaring integers; integers are always determined by the dimensions of the dataset
```

```
# for the current dataset x=1 (students) and y=62 (programs)
```

```
x = nrow(mydata)
```

```
y = ncol(mydata)
```

```
# declaring results matrix and aid matrix D
```

```
# result file has room for up to 40 variables, only 6 are used for the current application
```

```
results = matrix(nrow = x, ncol = 40)
```

```
D = matrix (nrow = y, ncol = 1)
```

```
# EASE algorithm for each student separately (cross-validation),
```

```
# indicated by the i in the for-loop
```

```
# this example only has one student (x = 1), the algorithm can run thousands of students  
# mainly depending on the processing power available  
# (cfr main paper, 7063 and 4892 students)  
# each different value for i will represent a different student  
# with a different model estimation, threshold, and advice set (size)
```

```
for (i in 1:x) {
```

```
# ordering correlations for one student from low to high
```

```
Dprep = Dprep[order(-Dprep[,i]),]
```

```
Dprep = as.matrix(Dprep)
```

```
# using an aid matrix D to calculate the parameters
```

```
# correlation (fit quality), correlation2, options and balance
```

```
D[,1] = Dprep[,i]
```

```
D = as.data.frame(D)
```

```
D$correlation = D$V1
```

```
D$correlation2 = D$correlation * D$correlation
```

```
D$options = seq(from = 1, to = y)
```

```
D$balance = D$correlation * D$options
```

```
# fitting quadratic (parabola) model  $ax^2 + bx + c$ 
```

```
fit=lm(balance ~ correlation2 + correlation, data = D)
```

```
summary(fit)
```

```
# plotting fit
```

```
par (cex = .8)
```

```
timevalues = seq(-1, 1, 0.01)
```

```
predictedcounts = predict(fit,list(correlation=timevalues, correlation2=timevalues^2))
```

```
plot(D$balance ~ D$correlation, col = "blue")
```

```
lines (timevalues, predictedcounts, col = "darkgreen",lwd = 3)
```

```
# extracting fit measure ( $R^2$ ) for the parabola model
```

```
results[i,1] = summary(fit)$r.squared
```

```
# extracting parameter weights a, b and c
```

```
coeff = tidy(summary(fit)$coefficients)
```

```
a = coeff[2,2]
```

```
b = coeff[3,2]
```

```
c = coeff[1,2]
```

```
# determining (ideal fit) threshold of the model  $ax^2 + bx + c$ 
```

```
ithreshold = (-1*b)/(2*a)
```

```
# extracting (ideal fit) threshold of the model  $ax^2 + bx + c$ 
```

```
results[i,2] = ithreshold
```

```
# extracting confidence interval ideal threshold
```

```
confint1 = confint_tidy(fit,conf.level=0.95)
```

```
# calculating the boundaries of the confidence interval (confintlow, confinhigh)
```

```
# using the confidence intervals of the parameter weights
```

```
confint2alow = -2*confint1[2,1]
```

```
confintblow = confint1[3,1]
```

```
confintlow = confint1[3,1] / (-2*confint1[2,1])
```

```
confinhigh = confint1[3,2] / (-2*confint1[2,2])
```

```
# extracting the boundaries of the confidence interval using the confidence intervals of
```

```
# the parameter weights
```

```
# confintlow corresponds to the actual threshold
```

```
results[i,3] = confintlow
```

```
results[i,6] = confinhigh
```

```
# testing uniformity of PE fit distribution
```

```
fit2 = lm(options ~ correlation, data=D)
```

```
summary(fit2)
```

```
# extracting estimation of uniformity of PE fit distribution
```

```
results[i,4] = summary(fit2)$r.squared
```

```
# how many choices does the student receive as part of his advice set?
```

```
results[i,5] = length( which( testmatrix[i,] > confintlow) )
```

```
}
```

```
# extract all results, adapt the path "C:/school/Revision paper 1 jee/" to where you want
```

```
# the "results" text file
```

```
# results include (from left to right)
```

```
# the fit of the EASE model (.996),
```

```
# the ideal threshold (.52),
```

```
# the actual threshold (.50),
```

```
# an estimate of the uniform distribution of PE fit in this specific student (.984
```

```
# size of the advice set (18)
```

```
# the upper boundary of the confidence interval of the ideal threshold (.55)
```

```
# 34 empty slots, placeholders for possible additional variables
```

```
write.table(results, "C:/school/Revision paper 1 jee/results", sep="\t")
```

end

7. Besides the results from the algorithm, the actual threshold (.50) can be cross-referenced to the list of programs and their PE fit (correlations) with the student profile to determine which programs exactly will be part of the advice set. In this case, all programs (18) with a PE fit over .50 are part of the advice set up to and including “industrial science: chemistry” (see below). The present example also indicates that the student has an interest profile corresponding to hard science programs, though not all hard science programs are part of the advice set. This process can be automatized further through use of an excel sheet or straight in the R-code if so desired.

<u>programs</u>	<u>PE fit</u>
biochemistry and biotechnology	.98
bioscience engineering: cell and gene biotechnology	.93
biology	.93
biomedical science	.89
physics and astronomy	.88
geology	.85
bioscience engineering: land and forest management	.84
mathematics	.83
veterinary medicine	.81
psychology: theoretical and experimental psychology	.81
chemistry	.80
pharmaceutical science	.80
bioscience engineering technology	.78
environmental engineering technology	.72
engineering: applied physics	.72
computer science	.67
bioscience engineering: environmental technology	.61
industrial science: chemistry	.55
geography	.49
dentistry	.47
archaeology	.45
bioscience engineering: agricultural science	.45

bioscience engineering	.44
medicine	.42
chemical engineering and materials science	.41
rehabilitation science and physiotherapy	.39
electrical engineering	.27
electromechanical engineering	.15
philosophy	.13
sociology	.13
physical education and movement science	.12
moral science	.10
African languages and cultures	.08
speech language and hearing science	.07
computer science engineering	.04
psychology: clinical psychology	.04
geomatics	.04
industrial science: electronics-ICT	.02
engineering: architecture	.01
industrial science: electro-mechanics	.01
information engineering technology	-.01
civil engineering	-.03
oriental languages and cultures	-.04
language and literature (two languages of choice like English and Dutch)	-.08
educational science: special education, disability studies and behavioral disorders	-.10
criminological science	-.11
East-European languages and cultures	-.13
art history, musicology and theatre studies	-.14
history	-.16
educational science: pedagogy	-.17
industrial design engineering technology	-.24
applied linguistics	-.24
civil engineering technology	-.30
economics	-.32
communication science	-.42
psychology: personnel management and industrial psychology	-.44
business engineering	-.45
political science	-.47
law	-.54
public administration and management	-.56
business economics	-.58
business administration	-.61

Chapter 3. The Influence of Interest Fit on (STEM) Study Choice: Female Students Fit their Choices Better.

Submitted to Research in Higher Education, 21st of May 2019.

Accepted for review, 23rd of May 2019.

*Stijn Schelfhout, Bart Wille, Lot Fonteyne, Elisabeth Roels, Filip De Fruyt &
Wouter Duyck (2019).*

3.1. Abstract

STEM (Science, Technology, Engineering and Mathematics) higher education study choice is considered the prime gateway towards the STEM field workforce, which is crucial to global economy. This gateway is recently threatened by two major problems. Indeed, STEM enrolments are declining while the STEM gender gap of female underrepresentation seems to widen. An important part of the solution for these problems may be to understand the underrepresentation of females in STEM programs from the perspective of vocational interests. However, literature almost exclusively focuses on gender differences regarding main interest dimensions, whereas the gender difference in the actual fit between students' interests and programs remains unexplored. This limits our understanding of the female underrepresentation in the STEM field, because literature already established that a good fit between students and programs predicts study choice, persistence and results. To address this issue, the present study investigates the effects of gender on study interests and interest fit with the chosen programs. Data was collected in the unique setting of an open access higher education system, with a particular focus on STEM study choice. More specifically, we assessed the interest fit of 9,162 first-year university students with their program of choice and with the STEM field. Results indicated that female students showed a better interest fit with their program than their male colleagues in both STEM and non-STEM programs. Moreover, STEM field interest fit contributed to the prediction of STEM study choice, and to an explanation of the STEM gender gap in enrolments. In order to promote STEM enrolment, orientation towards higher education should therefore also stress the interest fit of students with STEM programs. Special care should be taken regarding future female students, as their STEM choice seems to be even more linked to interest fit than for their male colleagues.

3.2. Introduction

STEM (Science, Technology, Engineering and Mathematics) study choice has become an important topic in vocational literature as it forms the primary gateway to the STEM work field (OECD, 2016; UNESCO, 2016). Research has shown that keeping this STEM work field well-staffed can be crucial to the economy of industrialized countries (World Economic Forum, 2016). However, this primary STEM gateway of higher education enrolment is threatened by two major problems. First, literature reports a decline in students choosing a STEM program in higher education (Ainley, Kos, & Nicholas, 2008; Perera & McIlveen, 2018). And second, literature also reports a widening gender gap in enrolments, indicating a growing female underrepresentation (Stoet & Geary, 2018; UNESCO, 2016; Xu, 2008). In order to investigate these issues, STEM field research has turned towards vocational interests, which are arguably considered to be the strongest predictors of study choice (Stoll et al., 2017; Wang & Degol, 2017). Indeed, recent meta-analyses on the STEM field workforce have repeatedly shown that STEM employees have high scores on specific single interest dimensions like the realistic or investigative dimension (Su & Rounds, 2015; Su, Rounds, & Armstrong, 2009). Males also seem to fit the STEM field better as they generally show higher interest in these dimensions, while females have higher scores on the social dimension. Interestingly, literature almost exclusively focuses on these gender differences on main interest dimensions, whereas no study assessed the gender difference in the actual fit between students' interests and programs. This limits our understanding of the female underrepresentation in the STEM field, because we already know that a good fit between students and programs has predictive validity towards study choice, persistence and results (Burns, 2014; Donnay, 1997; Nye, Su, Rounds, & Drasgow, 2012; Rounds & Su, 2014). Indeed, when assessing the interest fit between person and environment (PE interest fit), Holland (1997) always advocated the use of the full pattern of vocational interests, not only

the dominant dimensions. By matching the patterns of both interest profiles (student and program), the influence of interest fit on STEM study choice can be examined above and beyond the effects of individual interest dimensions. It is important to note that our study is conducted in an open access and low cost higher education system, where anyone with a high school degree can enroll for almost any program. Such an environment provides a unique opportunity to assess the influence of PE interest fit on study choice without risk of unwanted influence from high stakes testing or GPA requirements.

As such, the present study in an open access environment will address the presented issues in literature in two research questions. First, we want to investigate whether and to which extent male and female students differ regarding PE interest fit with the (STEM) program that they chose. Second, we want to investigate how interest fit contributes to the prediction of STEM study choice and an explanation of the STEM gender gap. With the answers to our questions, we aim to contribute to STEM and vocational interests literature so that researchers, policy makers and student counselors can act upon this knowledge to increase (female) student STEM enrolments.

The RIASEC Model of Vocational Interest and PE Interest Fit

Literature shows vocational interests predict up to 70 % of the study choice across students (Burns, 2014; Donnay, 1997; Paessler & Hell, 2012). Today, the RIASEC model by Holland (1997) is still one of the most influential models in vocational literature, describing the study interest profiles of students and programs through six RIASEC dimensions (realistic, investigative, artistic, social, enterprising and conventional). This model also displays an empirically verified circular structure: the dimensions are arranged in clockwise RIASEC order (Tracey & Rounds, 1995). To obtain a student's RIASEC profile, the literature describes a vast number of questionnaires, all rendering scores on the six dimensions (for an overview, see Nauta, 2010). The present research questions may be assessed with any

RIASEC questionnaire, but for the present study, we used SIMON-I, a validated instrument specifically targeting the transition from high school to higher education (Fonteyne, Wille et al., 2017). To obtain a program profile, student profiles can be used as representatives or *incumbents* for their program of choice. By averaging RIASEC scores of students enrolled in a program, a RIASEC profile can be established for each study program (Allen & Robbins, 2010). A recent study by Perera and McIlveen (2018) already showed that the individual dimensions of vocational interest were predictive of STEM study choice. Indeed, results suggested that students who score higher on the realistic, conventional, investigative, and enterprising dimensions chose STEM programs more often, with the size of these effects ranked accordingly from high to low.

The RIASEC model also allows for commensurate measurement of both the student's interest profile and the profile of the study program (Holland, 1997). In other words, both student and study program can be depicted on the same RIASEC circular structure through a score on all six dimensions. Such a commensurate measurement is key to determine the degree of person - environment interest fit (PE interest fit) between student and study programs. In other words, PE interest fit reflects how well a student matches his or her study program. Also, a higher PE fit has already been linked to higher autonomous and lower controlled motivation, indicating that a high PE fit might function as a motivator (Schelfhout et al., 2019). As an extra boon, the measure also uses the full two-dimensional circular structure of the RIASEC model, as all dimensions are accounted for.

STEM Study Choice and the Gender Gap

According to numbers from UNESCO (United Nations Educational, Scientific and Cultural Organization), female students only represent 35% of all students enrolled in higher education STEM programs and female researchers only account for 28% of all researchers active in the field (UNESCO, 2016). Despite efforts to get more women to the STEM field,

the gender gap keeps persisting. Other research seems to suggest that the STEM gender gap appears to be even larger in countries that advocate gender equality, and hence that perhaps growing gender equality may be related to the growing STEM gap. A large international study by Stoet and Geary (2018) showed that in developed, progressive and gender-aware countries, the need to choose STEM education for instrumental reasons (job prospects, salary, etc.) is smaller, so that women are more likely to choose non-STEM programs.

The gender gap in the STEM field seems to originate as early as primary school, and it is strongly tied to exposure to science and mathematics (Blackburn, 2017; Dejarnette, 2012). As such, girls are also underrepresented in STEM preparing-high school programs that focus on mathematics (Sadler, Sonnert, Hazari, & Tai, 2012; UNESCO, 2016). Moreover, Wang (2013a and 2013b) already showed that exposure to mathematics leads to an intent to major in STEM programs and thus holds predictive value towards higher education STEM choice. Though the overrepresentation of males is thus seemingly tied to more hours of high school mathematics, there is a specific proportion of female students that still enrolls for a STEM program. As such, female student STEM choice could be more tied to other variables like vocational interests.

As vocational interests are known to be strong predictors of study choice (Päßler & Hell, 2012), a specific and pronounced STEM interest could therefore be linked to why a specific group of female students still chooses a STEM program in higher education. A string of meta-analytic research on RIASEC gender differences in the more general STEM field workforce already revealed that males scored higher on the realistic and investigative dimensions (which are also typically tied to STEM professions) and lower on the artistic, social and conventional dimensions (Su, Rounds, & Armstrong, 2009). These gender differences were also confirmed to a large extent in research on a (non-STEM specific) open access study environment (Fonteyne, Wille et al., 2017). A recent follow-up on the meta-

analysis of Su and colleagues (2009) furthermore placed the gender differences on a *people / things* dimension on top of the RIASEC hexagon (Su & Rounds, 2015). Males had more interest in things (with high scores on realistic and investigative scales) while females were more interested in people (with high scores on the social scale). Su and Rounds (2015) also provided an explanation of the gender gap in the STEM field by placing work environments on the same *people / things* dimension. In sum, female students and female STEM employees thus seem to have lower scores on the typical STEM dimensions (R and I).

However, the presented literature does have some pending issues. First and foremost, research does not directly compare the STEM field to the non – STEM field. As a consequence, it also remains unknown how an interest fit with the STEM field contributes to STEM study choice and an explanation of the STEM gender gap. Second, the meta-analytic evidence did not study the influence of PE fit above and beyond the individual dimensions. And finally, the meta-analytic evidence targets the general STEM field workforce, not higher education as the primary gateway towards STEM employment. The present study will try and add to literature by resolving these issues.

Present Study

Considering PE interest fit predicts study choice (Burns, 2014; Donnay, 1997; Päßler & Hell, 2012), and considering the STEM choice in female students seems less tied to their hours of mathematics in high school (Blackburn, 2017; Dejarnette, 2012; Wang, 2013a & 2013b), a choice for the specific STEM field over a non-STEM career might be linked more strongly to PE fit in female students. As such, female students choosing STEM could display a very specific RIASEC profile, that possibly fits their program and the STEM field even better than their male colleagues. Our first research question will thus investigate whether and to which extent male and female student interest fit differs regarding their program of choice

and the STEM field. Tied to this research question are the following hypotheses that are tested within the STEM field,

H1: *female students have a higher interest fit with their STEM program of choice than male students.*

H2: *female students have a higher interest fit with the STEM field compared to male students.*

As PE fit predicts study choice, success and persistence (Burns, 2014; Donnay, 1997; Päßler & Hell, 2012; Nye et al., 2012; Rounds & Su, 2014), it is also crucial for future student orientation to know if this gender effect in STEM PE fit would generalize to a more broad population of students from STEM and non-STEM programs. Because we also have data on non - STEM programs, we have tested if the effect from H1 would generalize to the full student population,

H3: *female students have a higher interest fit with their program of choice compared to male students.*

Moreover, to be able to compare to the meta-analytic findings in the general work force, we also want to verify whether males also have an inherently better fit to the STEM field prior to employment (Su & Rounds, 2015; Su et al., 2009). As such, we have again tested the following hypothesis in the general student population,

H4: *male students have a higher interest fit with the STEM field compared to female students.*

For our second question, we have investigated how the STEM field interest fit contributes to predict STEM study choice and to explain the STEM gender gap. As such, we will also control for the already established effects in literature of gender (Su & Rounds, 2015), the RIASEC dimensions of vocational interest (Su et al., 2009; Fonteyne, Wille et al., 2017) and hours of mathematics in high school (Wille, Duyck et al., 2017, Pinxten et al.,

2017). We also explored how STEM field interest fit and the other main effects individually and uniquely contribute to the explained variance in STEM study choice. It is important to note that these effects have not been jointly tested in one model previously, making the outcome less predictable. As such, we have tested the following hypotheses,

H5: *STEM field interest fit predicts STEM study choice.*

H6: *the gender difference in STEM field interest fit contributes to a full explanation of the gender gap in STEM study choice.*

3.3. Method and Materials

STEM Field Definition and Application

UNESCO defines STEM as a field that incorporates traditional disciplines like mathematics and statistics, but also more specialized or technological disciplines like genetic engineering (UNESCO, 2016). The transition from high school to higher education is considered as a crucial timing to recruit future STEM field employees as it enrolls students from secondary education into a STEM program in higher education. For the present study, we have applied the UNESCO definition to student data collected in the context of the SIMON-project, which targets all faculties and all programs (STEM and non-STEM) at a Western European university (Shanghai Top 100, which ranks the world's top 1,500 universities and colleges based on objective measures). The data were primarily used to validate the online SIMON-instrument that focuses on the transition of individuals from high school to higher education by guiding these future students towards appropriate study programs, based on their skills (cognitive, non-cognitive and motivational factors; see Fonteyne, Duyck, & De Fruyt, 2017) and, as of 2016, also their vocational interests (see Fonteyne, Wille et al., 2017). Orientation for these students is necessary, as they are enrolling in open access and low cost (for the student) higher education, with nearly limitless possibilities. Indeed, in the present study's context, except for Medicine, Dentistry or

Performing Arts (music), all academic higher education programs are open to everyone with a high school degree. Moreover, the tuition fees do not exceed € 1,000 or about \$ 1,150 for one year, and almost half of the students receive funding through scholarships, based on economical (income) criteria. This open access environment provides a unique opportunity to assess the influence of study interests on study choice, without risk of unwanted influence from high stakes testing or GPA requirements.

Data and Procedure

During the start of the academic year 2016-2017 and 2017-2018, two cohorts of newly enrolled students filled out the online SIMON RIASEC interest questionnaires (September-October 2016 and 2017). Participation was not mandatory, but was promoted through professors, email and the online learning platform across all programs of the university, resulting in very high response rates. The questionnaires from the students enrolled for Medicine and Dentistry could not be considered for this investigation as they had to pass an exam to enroll for the program and thus formed the only exception on open access and study choice. The overall response rate amounted to 68% ($N = 9,162$, 60% female), with 3,389 students choosing a program in the STEM-field. Apart from these first-year students, we assessed the interest profiles of 39 study programs (see also Figure 1) using the interest RIASEC profiles of successful and persistent senior students, who indicated they would enroll again for the same program when given the opportunity ($N_0 = 6,572$). These senior students met the conditions of perseverance and academic success and the procedure of establishing the program E-profiles was identical to the procedure used by Allen and Robbins (2010). For each program, the RIASEC scores of all students were averaged for each dimension, resulting in a RIASEC profile for each program.

Measures

STEM or non-STEM study choice.

All 39 programs are divided into STEM and non-STEM programs based on the UNESCO definition (UNESCO, 2016). As a consequence, students in a (non-) STEM program are considered to have made a (non-) STEM study choice. Figure 1 shows a scatterplot of all programs. We can also make an approximation of the STEM field by averaging across all STEM programs. This calculation results in a STEM field RIASEC profile with the following dimension scores, R = 31.88, I = 46.25, A = 28.99, S = 24.54, E = 26.32 and C = 21.15.

Vocational interest

We used the SIMON-I questionnaire providing vocational interest scores on the six RIASEC dimensions (see APPENDIX A; Fonteyne, Wille et al., 2017). The RIASEC dimension scales obtained a reliability of .92, .88, .92, .92, .93 and .90 respectively (Cronbach's α). To test the assumed two-dimensional circular structure of the RIASEC dimensions, we first performed a confirmatory factor analysis (CFA), using the CirCe package in R (Browne, 1992; Grassi, Luccio, & Di Blas, 2010). The analysis confirmed the two-dimensional circular structure (Standardized Root Mean Square Residual of 0.053) and the parsimony (Schwarz's Bayesian Criterion of 0.03) of the RIASEC dimensions. Second, we also performed a *randomization test of hypothesized order relations* (RTOR) using the RANDALL package to confirm the circular structure and order of the RIASEC dimensions (Tracey, 1997). Results of this RTOR analysis revealed a *correspondence index* of .92, while the circular fit of the data also reached significance, $p = .02$. Both CFA and RTOR analysis thus indicated that our RIASEC data nicely fit a circular structure.

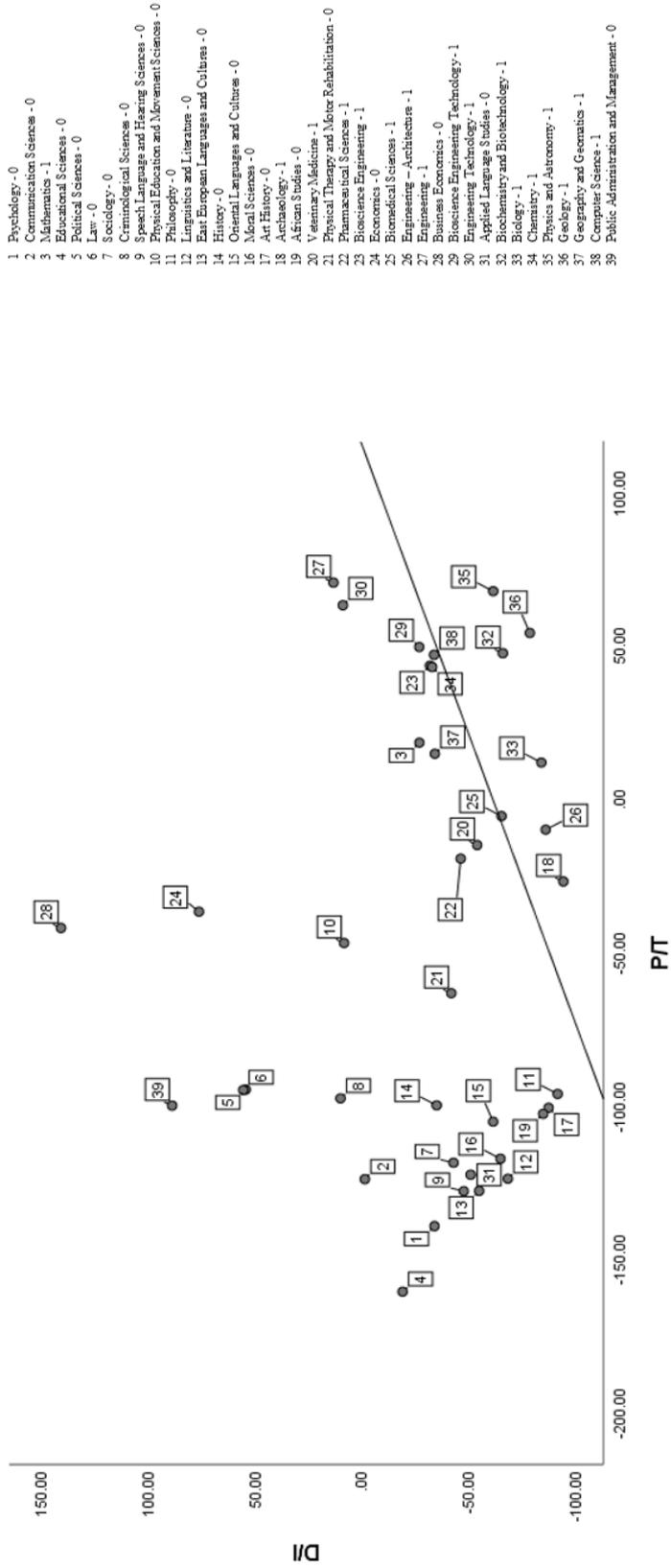


Figure 1. Scatterplot of 39 Programs Using the People / Things (X-axis) and Data / Ideas (Y-axis). STEM programs are annotated with a 1, non-STEM programs are annotated with a 0. All STEM programs are located in the right lower corner. The reference line indicates the relation between the P/T and D/I coordinates of the STEM programs,

$$y = 0.53x - 62.20$$

with an explained variance of 31%.

Program interest fit and STEM field interest fit

We measured PE interest fit using Euclidean distance as described by Wille, Tracey, Feys and De Fruyt (2014). The elaboration of Holland's RIASEC model with the so-called Prediger dimensions of *People / Things* (P/T) and *Data / Ideas* (D/I) enables the use of Euclidean distance as a two-dimensional PE interest fit measure (Prediger, 1982; Prediger, 2000). Practically, Euclidean distance can be calculated using the following formulae (retrieved from Wille and colleagues, 2014): $P / T = 2 \times R + I - A - 2 \times S - E + C$, $D / I = (1.73 \times E + 1.73 \times C - 1.73 \times I - 1.73 \times A)$ and Euclidean distance = $\text{SQRT} ((\text{student } P / T - \text{study program } P / T)^2 + (\text{student } D / I - \text{study program } D / I)^2)$. A low distance between student and program profiles indicates a better PE interest fit. Studies have shown that a low Euclidean distance holds predictive validity towards degree attainment (Tracey, Allen, & Robbins, 2012). Degree attainment is important for the present study as it provides access to the STEM work field. Utilizing the established RIASEC profiles for students and programs, we thus calculated the Euclidean distance for each student with his or her chosen program. As we need an interest fit measure towards the STEM field for our STEM study choice model, we also calculated the Euclidean distance between students and the STEM field (with coordinates $P/T = 26.76$ and $D/I = -48.04$). To avoid confusion, program interest fit indicates the fit between a student and her/his program, while STEM field interest fit indicates the fit between a student and the STEM field. A lower Euclidean distance indicates a better fit for both variables.

Hours of high school mathematics

For the present study, weekly hours of high school mathematics prior to higher education enrolment will act as a control for the influence of STEM field interest fit on STEM study choice.

Analyses

To test **H1** to **H4**, we used a Welch two-sample, two-tailed t-test. Effect sizes are calculated using a Cohen's *d* (Sawilowsky, 2009). **H1** and **H2** are tested in a STEM specific student population, **H3** to **H4** are tested in the general population. Apart from these hypotheses, we also analyzed the gender differences in all RIASEC dimensions and the hours of high school mathematics to be able to integrate our findings into literature. To test **H5** and **H6**, we constructed a STEM study choice model, which is a logistic regression of STEM study choice on the main effects and gender interactions of the hours of high school mathematics, PE interest fit and all six base RIASEC interest dimensions. Though the logistic model is tailored towards STEM study choice, it also directly implies a profile towards non-STEM study choice as the outcome is binary (1 or 0). As STEM is the focus of this study, we formulated the results in terms of STEM study choice. The model was built in two stages by first adding all main effects, followed by all student gender interactions. As our distinctive model has to deal with a fairly large amount of predictors by adding the RIASEC dimensions and all possible gender interactions, we used Akaike's Information Criterion (AIC) in a stepwise selection procedure to select the best fitting model, distinguishing students that chose STEM from those that chose another (non-STEM) program. More formally,

$$AIC = wk - 2 \ln L(m)$$

with *k* representing the number of estimated parameters, *w* representing the weight of the parameter term and *L(m)* representing the maximum likelihood for model *m* (Burnham & Anderson, 2002). From a set of all possible models with all possible predictors, the stepwise procedure selected the best fitting one with the lowest AIC. This procedure rewards models with the least chance of information loss, but penalizes models that use too many predictors. The AIC stepwise methodology has a number of advantages over classic stepwise regression. AIC does not use statistical testing as a criterion for model selection and does not depend on when variables enter the equation as all possible models are considered. After selecting the

best fitting model, we performed a logistic regression with STEM choice as dependent variable and the variables from the selected model as predictors. We also reported two additional measures of explained variance (deviance) concerning the individual main effects to estimate their specific contribution towards STEM choice prediction and as control variables for the effects of interest fit. First, individual explained variance indicates how much variance the predictor explains if there are no other predictors present in the model. Second, unique explained variance indicates how much explained variance is lost if the predictor is removed from the model. To conclude, we constructed a ROC curve (receiver operating characteristic curve) indicating how well our model succeeds in profiling STEM students and distinguishing them from their non-STEM colleagues. A ROC curve always balances sensitivity and specificity. Sensitivity indicates the proportion of STEM students that were actually classified as STEM students by our STEM choice model, while specificity indicates the non-STEM students that were indeed classified as non-STEM students. A rising sensitivity results in a falling specificity and vice versa. Finally, the area under the curve (AUC) indicates how well the model can make the distinction between STEM and non-STEM students.

3.4. Results

Regarding our first question, Tables 1 (**H1** and **H2**) and 2 (**H3** and **H4**) show the descriptive statistics and the gender differences for program interest fit and STEM field interest fit, the base RIASEC dimension scales and the hours of high school mathematics. For **H1** and **H2**, our predictions were confirmed: female students in the STEM field interest fit their program of choice and the STEM field better (lower Euclidean distance) than their male colleagues. For **H3**, we observe that female students also fit their program better in the overall student population, generalizing the effect from the STEM specific population as expected.

Finally for **H4**, we also observe that the male students in the general student population have a better fit than their female colleagues, again confirming our hypothesis.

Table 1. Student Gender Differences in the STEM Student Population.

Predictors	<i>n</i>	<i>M</i>	<i>M_m</i>	<i>M_f</i>	<i>SD</i>	Cohen's <i>d</i>
program interest fit	3,389	83.15	85.95	79.88	50.26	0.12
STEM field interest fit	3,389	93.06	97.75	87.57	53.73	0.19
realistic dimension	3,389	33.79	45.94	19.56	26.96	1.13
investigative dimension	3,389	44.88	41.94	48.33	20.96	-0.31
artistic dimension	3,389	23.38	20.46	26.80	22.52	-0.28
social dimension	3,389	22.67	15.21	31.40	21.50	-0.80
enterprising dimension	3,389	23.32	25.95	20.23	22.47	0.26
conventional dimension	3,389	17.49	18.49	16.32	19.24	0.11
hours of high school mathematics	3,377	6.11	6.42	5.75	1.51	0.45

Note. M_m = male student average and M_f = female student average. Hours of high school mathematics was operationalized through the hours of mathematics students chose in the final two years of high school up to a maximum of eight. The RIASEC dimensions were measured on a scale from 1 to 100. A positive student gender difference indicates higher male student scores, a negative difference indicates higher female student scores. Cohen's *d* effect size rules of thumb (Sawilowski, 2009): 0.01 – very small effect, 0.20 – small effect, 0.50 – medium effect, 0.80 – large effect, 1.20 – very large effect, 2.00 – huge effect. All gender differences were significant at the level $p < .001$.

Table 2. Student Gender Differences in the General Student Population.

Predictors	<i>N</i>	<i>M</i>	<i>M_m</i>	<i>M_f</i>	<i>SD</i>	Cohen's <i>d</i>
program interest fit	9,162	85.39	90.37	82.00	48.35	0.17
STEM field interest fit	9,162	138.77	126.55	147.08	72.44	-0.29
realistic dimension	9,162	18.86	32.00	9.93	23.61	1.00
investigative dimension	9,162	33.50	34.48	32.83	21.29	0.08
artistic dimension	9,162	29.95	24.62	33.57	25.59	-0.36
social dimension	9,162	34.93	22.70	43.23	25.90	-0.88
enterprising dimension	9,162	33.45	37.23	30.87	28.17	0.23
conventional dimension	9,162	21.08	25.08	18.36	22.86	0.29
hours of high school mathematics	9,135	4.95	5.41	4.59	2.88	0.47

Note. Hours of high school mathematics was operationalized through the hours of mathematics students chose in the final two years of high school up to a maximum of eight. The RIASEC dimensions were measured on a scale from 1 to 100. A positive student gender difference indicates higher male student scores, a negative difference indicates higher female student scores. Cohen’s *d* effect size rules of thumb (Sawilowski, 2009): 0.01 – very small effect, 0.20 – small effect, 0.50 – medium effect, 0.80 – large effect, 1.20 – very large effect, 2.00 – huge effect. All gender differences were significant at the level $p < .001$.

To start the analyses of our second question, Table 3 shows the overall student gender proportions versus the STEM specific student gender proportions for the current sample. The chi-square test on these proportions was significant, $\chi^2(1) = 405.62, p < .001$, rejecting the null hypothesis and indicating a gender gap does exist regarding STEM study choice in the current data sample. The gender gap is characterized through a male overrepresentation in STEM study choice.

Table 3. Student Gender and STEM Choice Cross-Tabulation.

		STEM choice		total
		0	1	
male	N	1,879	1,828	3,707
	% within student gender	51	49	100
	% within STEM choice	33	54	40
female	N	3,894	1,561	5,455
	% within student gender	71	29	100
	% within STEM choice	67	46	60
total	N	5,773	3,389	9,162

Note. STEM = Science, Technology, Engineering, Mathematics.

Next, we performed a logistic regression on the final STEM study choice model after the AIC procedure. Three student gender interactions (student gender \times R, student gender \times A and student gender \times E) were removed by the AIC stepwise regression as they did not add to the prediction of STEM study choice. Table 4 shows the main- and interaction effects of the final model, explaining about 71% of the (pseudo-) variance through a Nagelkerke’s R^2 . The

student gender main effect does not reach significance ($p = .15$), indicating that the student gender difference in making a STEM study choice is fully explained through the student gender interactions present. Table 5 compares the individual and unique explained variance of STEM field interest fit to the other predictors in the model. All predictors have a somewhat low and thus similar unique explained variance. This indicates that the model is quite robust. Indeed, information loss remains limited when removing one predictor from the model. In contrast, the individual explained variance over all predictors shows a much wider range. Important to note, the variance measures for student gender provide an additional indication of the gender gap. About one to six percent of STEM study choice can be explained through student gender (without gender interactions), while controlling for the other variables.

Table 4. STEM Study Choice Model: Coefficients.

coefficients	estimate	z - statistic
(Intercept)	-3.13	-13.38***
student gender	0.44	1.43
STEM field interest fit	-0.0068	-5.65***
hours of high school mathematics	0.59	16.78***
realistic dimension	0.049	22.39***
investigative dimension	0.039	12.70***
artistic dimension	-0.014	-7.60***
social dimension	-0.038	-11.26***
enterprising dimension	-0.013	-6.13***
conventional dimension	-0.028	-8.09***
student gender \times STEM field interest fit	-0.0053	-3.19**
student gender \times hours of high school mathematics	-0.14	-3.71**
student gender \times investigative dimension	0.022	5.28***
student gender \times social dimension	0.015	3.29**
student gender \times conventional dimension	0.021	5.03***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

Table 5. STEM Study Choice Model: Individual and Unique Explained Variance of Predictors.

predictors	individual explained variance	unique explained variance
student gender	0.06	0.01
STEM field interest fit	0.32	0.01
hours of high school mathematics	0.34	0.05

realistic dimension	0.30	0.05
investigative dimension	0.22	0.05
artistic dimension	0.05	0.01
social dimension	0.19	0.02
enterprising dimension	0.11	< 0.01
conventional dimension	0.02	< 0.01

Note. Individual and unique explained variance were measured using a Nagelkerke's R^2 .

Regarding **H5**, we observe that STEM field interest fit has a significant effect on STEM study choice, even when controlling for gender, hours of high school mathematics and the base RIASEC dimensions. Students with a higher STEM field interest fit have a higher chance of choosing a STEM study program. STEM field interest fit also shows a high individual variance, second only to hours of high school mathematics. Given these results, our hypothesis was confirmed: STEM field fit does predict STEM study choice.

Regarding **H6**, we observe that female students with a good STEM field interest fit have a higher chance of choosing a STEM program compared to their male colleagues (interaction effect). In other words, the positive effect of a good PE interest fit on STEM study choice is amplified for female students. These findings thus confirm our initial hypothesis. Important to note, the main effect and the gender interaction of the hours of high school mathematics were also significant. Students with more hours of high school mathematics have a higher chance of choosing STEM, an effect that is even stronger for males. This confirms our assumption from the introduction that student STEM choice is indeed more strongly related to hours of high school mathematics in male students than in female students (Blackburn, 2017; Dejarnette, 2012; Wang, 2013a & 2013b).

As the validity of these results is highly dependent on the ability of our STEM study choice model to distinguish STEM and non-STEM students, we plotted the balance between sensitivity and specificity of the model on a ROC curve in Figure 2. As an example, our STEM study choice model, only using the mentioned interest variables, succeeds in correctly identify 87% of the students that indeed chose a STEM program (sensitivity), while it also

manages to correctly identify 87% of the students that chose a non-STEM program (specificity). Finally, analyses also revealed an AUC of 0.94 with an asymptotic 95% CI of [0.938, 0.947], indicating an excellent fit.

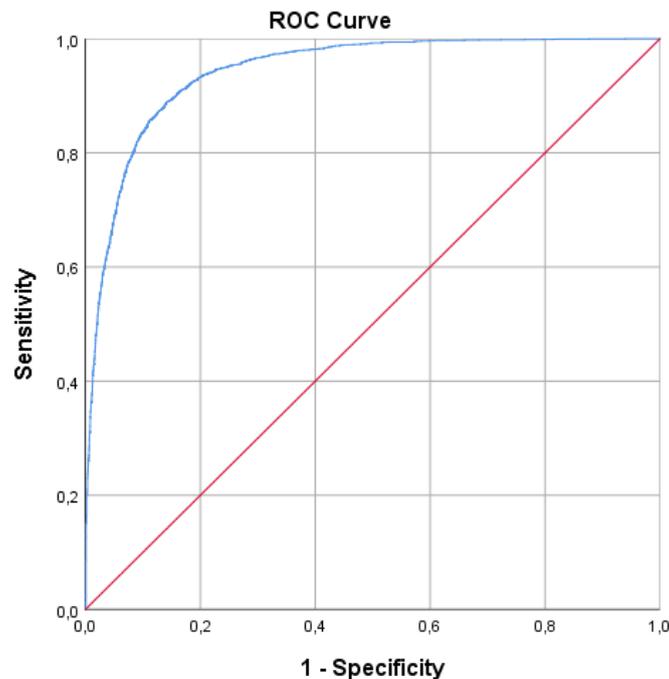


Figure 2. ROC Curve (in blue) of (non-) STEM Study Choice Distinction.

Specificity is displayed on the X-axis and sensitivity on the Y-axis. The 50% reference line is indicated in red.

3.5. Discussion

Industrialized regions around the globe have experienced increasing difficulty to fill out vacant STEM positions due to a decline in students who actively enroll for a STEM program in higher education (Ainley et al., 2008; Perera & McIlveen, 2018). Also, there seems to exist a (widening) STEM gender gap, indicating that women are underrepresented in the STEM field (UNESCO, 2016; Xu, 2008). To ensure a steady stream of (female) students into higher education, literature needs to identify determinants of STEM study choice and formulate an explanation towards the gender gap, so that researchers, student counselors and policy makers can act upon this knowledge to attract more (female) students. In this context, the present study presented two research questions in an academic higher education

environment. For our first research question, we have investigated whether and to which extent male and female students differ regarding PE interest fit with the higher education programs that they choose. For our second question, we have investigated how interest fit contributes to prediction of STEM study choice and explanation of the STEM gender gap. It is important to note that our study was conducted in an open access academic higher education environment, which allows an assessment of the effects of interest (fit) without the external constraints that are imposed in systems that use high stakes tests or entrance requirements.

Our first question revealed that the average program interest fit of female students with their chosen STEM study program was on average about 7% better in comparison to their male colleagues². Female students in the STEM field also had a 10% better fit with the STEM field compared to their male colleagues. As higher education is a crucial stream towards the STEM workforce, this is an important result (UNESCO, 2016). These effects of increased PE fit for female students also generalized to the full student population, with female students now showing a program fit that was on average 9% better than their male colleagues. To our knowledge, a two-dimensional PE interest fit gender difference has not yet been reported in (STEM) literature. As an interpretation of the result, we speculate that the open access environment reflects pure interest in the study program or the STEM field, as there are no access requirements like entry exams. The results thus suggest that (STEM) study choice in female students is linked more strongly to PE interest fit and perhaps less to instrumental motives (Schelfhout et al., 2019). The results could also have serious ramifications for general student orientation advice as PE fit predicts study choice, success and persistence (Burns, 2014; Donnay, 1997; Päßler & Hell, 2012; Nye et al., 2012; Rounds

² comparing Euclidean distance, 79.88 (M_f) divided by 85.95 (M_m) is 0.93 or 7% lower.

& Su, 2014). It would also be very interesting to see how this finding generalizes to closed access study environments. Furthermore, male students had a 14% better average interest fit with the STEM field compared to their female colleagues for the general student population, prior to employment. These findings expand meta-analytic workforce research on the STEM-field (Su & Rounds, 2015). We have also found that male students scored higher on the realistic, investigative, enterprising and conventional RIASEC dimensions, while female students scored higher on the artistic and social dimensions. In the STEM specific student population, results were analogous, barring one exception. Female students now scored 15% higher on the investigative dimension scale compared to their male colleagues, reversing the effect from the general student population. The results diverge from literature on the STEM workforce (Su et al., 2009), but they are consistent with earlier findings from higher education (Fonteyne et al., 2017). Male students also had 18% (12% in the STEM specific population) more hours high school mathematics compared to their female colleagues. This result is completely in line with literature (Sadler et al., 2012; UNESCO, 2016; Wang 2013a and 2013b).

Our second question assessed how STEM field interest fit contributes to the prediction of STEM choice and an explanation of the gender gap, while controlling for gender, the RIASEC dimension scores, and hours of high school mathematics. We started with the observation that there was in fact a STEM gender gap in our data, characterized by a male overrepresentation. Indeed, only a mere 29% of all female students made a STEM study choice, in contrast to about half of the male students. The gender factor was responsible for up to 6% of student (non-) STEM choice. As a result, the STEM study field consisted of 46% female students. However, our study did include 60% female students to begin with, somewhat creating a better gender balance compared to literature (UNESCO, 2017; Xu, 2008). Taken together, the small proportion of female students choosing STEM seems to be in

line with the results from the broader STEM gender gap presented by Stoet and Geary (2018) for developed, progressive and gender-aware countries. As a possible explanation, our open access higher education environment in a developed, progressive and gender-aware country invites students to choose according to their interests. As such, the present study already confirmed the findings in literature that males have a higher interest in the STEM field (Su et al., 2009; Su & Rounds, 2015; UNESCO, 2016).

Next, we constructed a STEM study choice model integrating the effects of student gender and its interactions, the RIASEC dimensions, STEM field interest fit and hours of high school mathematics. The model succeeded in fully explaining the gender gap and performed adequately, with explaining up to 71% of the variance in (non-) STEM study choice. The model also succeeded in profiling STEM and non-STEM students by correctly identifying their study choice in 87% of all cases. These numbers are on the very high end when compared to known (vocational) literature on study choice (Burns, 2014; Donnay, 1997; Päßler & Hell, 2012). Moreover, the model seemed quite robust against information loss. Indeed, all predictors had low unique explained variance, ranging from (less than) 1 to 5%. As such, little information is lost when a single predictor is removed.

Most importantly, we have found that STEM field fit was predictive of STEM study choice, even when controlling for the effects of the base RIASEC dimensions and hours of high school mathematics. Students with a better PE fit (lower Euclidean distance) thus have a higher chance of choosing STEM. Moreover, this effect was even stronger in female students. To our knowledge this influence of STEM field fit on STEM study choice and on the STEM gender gap has not yet been reported in (STEM) literature. This confirms that, in order to understand female underrepresentation in STEM, it is essential to not only look at gender effects on interest dimension, but also to include gender differences in interest fit. Interestingly, our study showed that the hours of high school mathematics has a positive

effect towards STEM choice, especially in male students. This find is completely in line with literature on STEM preparation (Blackburn, 2017; Dejarnette, 2012; Wang 2013a and 2013b).

The RIASEC effects found are somewhat in line with STEM field literature (Su et al., 2009; Su & Rounds, 2015). Students with a high realistic and investigative dimension indeed have a higher chance of a STEM study choice. However, our study also showed that all interest dimensions are important as students with high artistic, social, enterprising and conventional interest dimensions have a lower chance of a STEM study choice. These findings thus indicate that all RIASEC dimensions should therefore be taken into consideration for each data sample, although the pattern of the individual effects might vary depending on the nature of the sample (Holland, 1997; Perera & McIlveen, 2018). Apart from these main effects, a high investigative dimension has an even larger positive influence in female students, while a high social and conventional dimension have a less negative influence. These specific RIASEC patterns are largely at odds with literature on the STEM field (Su et al., 2009; Su & Rounds, 2015). As an explanation, we suggest that the open access higher education environment allows especially female students to choose a program that matches their interests.

Though the present study clearly shows STEM study choice can be almost fully understood through a study interest perspective, we did not assess and therefore are neutral to the question what the origin of these interest differences might be. Indeed, vocational interest gender differences in students aged 18 does not rule out nor confirms an explanation featuring genetics, socialization or a combination of both. As such an explanation is beyond the scope of our present study, we will not discuss this issue any further at this moment.

Conclusion

In an open access study environment, female students fit their study choice better than their male colleagues. Alongside the classic RIASEC interest dimensions and hours of high

school mathematics, interest fit contributes to specific profiles that predict STEM study choice and explain the STEM gender gap. In order to promote STEM enrolment, orientation towards higher education should therefore also stress the fit of students with STEM programs. In order to bridge the STEM field gender gap, special care should be taken regarding future female students, as their STEM choice seems to be even more linked to interest fit.

3.6. References

- Ainley, J., Kos, J., & Nicholas, M. (2008). Participation in Science, Mathematics and Technology in Australian Education. *ACER Research Monographs*. Retrieved 11th of June 2019, from http://research.acer.edu.au/acer_monographs/4
- Blackburn, H. (2017). The Status of Women in STEM in Higher Education: A Review of the Literature 2007–2017. *Science and Technology libraries*, 36:3. doi: 10.1080/0194262X.2017.1371658
- Browne, M. W. (1992). Circumplex models for correlation matrices. *Psychometrika*, 57(47), 469-497. doi:10.1007/bf02294416
- Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A practical information-theoretic approach* (2nd ed.), Springer-Verlag.
- Burns, S. T. (2014). Validity of Person Matching in Vocational Interest Inventories. *The Career Development Quarterly*, 62(2), 114-127. doi: 10.1002/j.2161-0045.2014.00074.x
- Dejarnette, N. (2012). America's children: providing early exposure to STEM (Science, Technology, Engineering and Math) initiatives. *Education*, 133(1), 77–84.
- Donnay, D. (1997). E. K. Strong's legacy and beyond: 70 years of the Strong Interest Inventory. *The Career Development Quarterly*, 46(1), 2–22. doi: 10.1002/j.2161-0045.1997.tb00688.x
- 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. doi: 10.1016/j.lindif.2017.05.003
- Fonteyne, L., Wille, B., Duyck, W., & De Fruyt, F. (2017). 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*, 17(2), 233-262. doi:10.1007/s10775-016-9327-9
- Grassi, M., Luccio, R., & Di Blas, L. (2010). CircE: An R implementation of Browne's circular stochastic process model. *Behavior Research Methods*, 42(1), 55–73. doi: 10.3758/BRM.42.1.55
- Holland, J.L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, FL: Psychological Assessment Resources.
- 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 (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. PISA, OECD Publishing, Paris. doi: 10.1787/9789264266490-en.
- Päßler, K., & Hell, B. (2012). Do interests and cognitive abilities help explain college major choice equally well for women and men? *Journal of Career Assessment*, 20(4), 479–496. doi: 10.1177/1069072712450009
- Perera, H. N., & McIlveen, P. (2018). Vocational interest profiles: Profile replicability and relations with the STEM major choice and the Big-Five. *Journal of Vocational Behavior*, 106, 84-100. doi: 10.1016/j.jvb.2017.11.012
- Pinxten, M., Van Soom, C., Peeters, C., De Laet, T., & Langie, G. (2017). At-risk at the gate: prediction of study success of first-year science and engineering students in an open-admission university in Flanders—any incremental validity of study strategies?

- European Journal of Psychology of Education*, 34(1), 45-66. doi: 10.1007/s10212-017-0361-x
- Prediger, D. J. (1982). Dimensions underlying Holland's hexagon: Missing link between interests and occupations? *Journal of Vocational Behavior*, 21(3), 259-287.
- Prediger, D. J. (2000). Holland's hexagon is alive and well—though somewhat out of shape: Response to Tinsley. *Journal of Vocational Behavior*, 56(2), 197–204.
- Rounds, J., & Su, R. (2014). The Nature and power of Interests. *Current Directions in Psychological Science*, 23(2), 98-103. doi: 10.1177/0963721414522812
- Sadler, P. M., Sonnert, G., Hazari, Z., & Tai, R. (2012). Stability and Volatility of STEM Career Interest in High School: A Gender Study. *Science Education*, 96(3), 411–427. doi: 10.1002/sce.21007
- Sawilowsky, S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, 8(2), 467-474.
- Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., De Fruyt, F., & Duyck, W. (2019). The effects of vocational interest on study results: Student person – environment fit and program interest diversity. *PLoS ONE*. doi: 10.1371/journal.pone.0214618
- Stoet, G., & Geary, D. C. (2018). The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education. *Psychological Science*, 29(4), 581-593. doi: 10.1177/0956797617741719
- Stoll, G., Rieger, S., Lüdtke, O., Nagengast, B., Trautwein, U., & Roberts, B.W. (2017). Vocational interests assessed at the end of high school predict life outcomes assessed 10 years later over and above IQ and Big Five personality traits. *Journal of Personality and Social Psychology*, 113(1), 167-184. doi: 10.1037/pspp0000117

- Su, R., & Rounds, J. (2015). All STEM fields are not created equal: People and things interests explain gender disparities across STEM fields. *Frontiers in Psychology*, 6:189. doi: 10.3389/fpsyg.2015.00189
- 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
- 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. (1995). The Arbitrary Nature of Holland's RIASEC Types: a Concentric-Circles Structure. *Journal of Counseling Psychology*, 42(4), 431-439. doi: 10.1037/0022-0167.42.4.431
- 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.
- UNESCO (2016). Sharing Malaysian Experience in Participation of Girls in STEM Education, *Current and Critical Issues in Curriculum and Learning*, retrieved 3rd October 2018, from <https://unevoc.unesco.org/go.php?q=TVETipedia+Glossary+A-Z&id=640>
- Wang, M.T. & Degol, J. L. (2017). Gender Gap in Science, Technology, Engineering, and Mathematics (STEM): Current Knowledge, Implications for Practice, Policy, and Future Directions. *Educational Psychology Review*, 29(1) 119-140. <https://doi.org/10.1007/s10648-015-9355-x>

- Wang, X. (2013a). Modeling Entrance into STEM Fields of Study Among Students Beginning at Community Colleges and Four-Year Institutions. *Research in Higher Education*, 54(6), 664-692. doi: 10.1007/s11162-013-9291-x
- Wang, X. (2013b). Why Students Choose STEM Majors: Motivation, High School Learning, and Postsecondary Context of Support. *American Educational Research Journal*, 50(5), 1081-1121. doi: 10.3102/0002831213488622
- 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
- World Economic Forum (2016). *The Human Capital Report*, retrieved 2 October 2018, from <http://reports.weforum.org/human-capital-report-2016/>
- Xu, Y. J. (2008). Gender Disparity in STEM Disciplines: A Study of Faculty Attrition and Turnover Intentions. *Research in Higher Education*, 49(7), 607-624.
doi:10.1007/s11162-008-9097-4

3.7. APPENDIX A:

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.

Dimensions were masked for the participant.

	Dimension	YES	NO
Developing electronic systems	R		
Analysing the grammatical structure of a sentence	I		
Helping people with speech disorders	S		
Recruiting a job candidate	E		
Monitoring the quality standards for food safety and hygiene	C		
Repairing malfunctioning electrical equipment	R		
Carrying out laboratorial analyses	I		
Designing a poster for an exhibition	A		
Helping others with their personal problems	S		
Organising a conference	E		
Preparing financial reports	C		
Being responsible for the maintenance of IT hardware	R		
Analysing statistics	I		
Designing webpages	A		
Developing council prevention campaigns	S		
Presenting new policy propositions	E		
Collecting quantitative and qualitative data	I		
Develop new methods for industrial production	R		
Treating diseases in animals	I		
Editing the sound and images for a movie	A		
Formulating education and training policies	S		
Drawing up the budgets	C		
Doing the follow up on building sites	R		
Analysing x-rays/brain scans	I		
Fit out a show room	A		
Sport guidance for children, the elderly, ...	S		
Formulate a theory about the differences between population groups	I		
Monitor quality standards	C		
Maintaining airplanes	R		
Investigating the impact of historical people	A		
Composing a work of music	A		
Providing guidance for victims	S		
Selling a product or service	E		
Calculating prices	C		
Installing and maintaining computer servers	R		
Designing an advertising folder	A		
Providing information about the assistance for the poor	S		

Drawing up an organisational business or policy plan	E
Checking bank transactions	C
Developing windmill parks	R
Prove a theorem	I
Analysing text structures	A
Giving travel advice	S
Negotiating contracts	E
Drawing up a contract	C
Investigating chromosomal defects	I
Writing scenarios	A
Holding tests, questionnaires and in-depth interviews	S
Screening the administration	C
Working on a drilling rig	R
Turning an idea into a film	A
Giving care to patients	S
Restructuring an organisation or company	E
Checking the compliance of regulations	C
Excluding alternative explanations through experiments	I
Designing the layout of a hospital	A
Advising youngsters regarding their vocational choice	S
Exploring new economic markets	E
Drawing up the annual report	C
Setting up a festival stage	R
Developing a new medicine	I
Writing a review	A
Giving training in communication skills	S
Starting up an enterprise	E
Investigating a cost structure	C
Creating a technical drawing	R
Putting theories in their historical and social context	I
Creating an art piece	A
Giving health advice	S
Giving health and parenting education	E
Calculating expenses	C
Disassembling electrical appliances	R
Comparing cultures	A
Guiding minority groups on the job market	S
Conducting a meeting	E
Drawing up a timetable	C
Measuring a lane	R
Supporting and following up foster families	S
Attracting sponsors	E
Standing in front of a classroom	S
Leading a team	E
Managing a database	C
Collecting soil samples	R

Beginning a herbarium (a plant collection)	I
Counseling underprivileged people	S
Formulating a treatment plan	S
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.

Dimensions were masked for the participant.

	Dimension	YES	NO
Industrial designer	R		
Civil engineer	I		
Fashion designer	A		
Policy advisor in political and international relations	E		
Recruitment and selection advisor	E		
Damage expert	C		
Agricultural technician	R		
Teacher	S		
Business economist	C		
Accountant	C		
Electrical engineer	R		
Biologist	I		
Art/music teacher	A		
Speech therapist	S		
Bank manager	C		
Landscape architect	R		
Physicist	I		
Editor	A		
Student counselor	S		
Tax supervisor	C		
Neurologist	I		
Policy advisor art and culture	A		
Educator	S		
Marketing manager	E		
Safety advisor	C		
Construction manager	R		
Historian	I		
Director	A		
Communication manager	E		
Manager (of a company)	E		
Judge	C		
Forester	R		
Researcher	I		
Graphic designer	A		
Psychologist	S		

Lawyer	E
Notary	C
Mathematician	I
Art historian	A
Social worker	S
Politician	E
Pilot	R
Pharmacist	I
Linguist	A
Divorce mediator	S
Journalist	A
Structural engineer	R
Lab assistant	I
Photographer	A
Nurse	S
Advertising campaign manager	E
Chemist	I
Tax specialist	C
Architect	R
Artist	A
Educational scientist	S
Librarian	A
Philosopher	I
Representative	E
Geneticist	I
Interior designer	A
Estate agent	E
Physiotherapist	S
Meteorologist	I
Sales manager	E
Statistician	I

Chapter 4. Identifying Students at Risk of Failing in Academic Higher Education.

Submitted to Learning and Individual Differences, 27th of May 2019.

Accepted for review, 11th of June 2019.

*Stijn Schelfhout, Bart Wille, Lot Fonteyne, Elisabeth Roels, Filip De Fruyt &
Wouter Duyck (2019).*

4.1. Abstract

In many Western-European countries, access to higher education is relatively unconstrained. After enrolment, students have to make a hefty investment in time and resources to obtain the desired degree. Typically however, such systems have low first-year pass rates because a lot of students do not seem to possess the basic (cognitive) skills, interest or motivation to succeed. By identifying such students at risk early on, academic institutions may offer advice and remediation opportunities, or reorient them towards more suitable programs. As such, the present study wants to answer the question how well we can identify failing students (given the focus on not passing, *true positives*), without inflating the number of successful students identified as failing (*false positives*). Up to now, educational and psychological literature almost exclusively focuses on the explanation of population variance in academic achievement, often measured by grade point average (GPA). Here, we will instead focus on a predictive and program-specific model that combines known explanative predictors of the variance in academic achievement in order to validate them for individual student prediction ($N = 6,624$) of failing or passing. As such, at a 5% false positive rate in an open access academic environment, our study shows we can accurately identify 29% of all students at risk of failing, with a program-specific maximum of 66%. More lenient false positive rates enable to identify 58% of all students at risk, with a program-maximum of 81%. Study antecedents and cognitive ability predict student academic achievement in a wide range of study programs, followed by less frequent predictors such as motivation, vocational interest, cognitive test anxiety, academic self-efficacy, self-control, conscientiousness and metacognition. These findings are crucial in providing practical advice to students at risk of failing.

4.2. Introduction

Higher education is a taxing endeavor as students have to make a significant investment of time and resources. Especially in many Western-European open access systems, where everyone with a high school degree can enter almost any program, knowledge about avoiding study failure would help in identifying those students who do not have the basic skills (like cognitive ability) or properties (like interest in a study program) to complete the curriculum. By identifying these students early on, an advice towards extra schooling may remedy the student's basic skills or an advice towards reorientation can relocate these students towards a more appropriate program. This way, both students and universities can actually save time, resources and effort, while success rates for individual programs will improve. One such tool from a Western university (Shanghai top 100, www.shanghairanking.com), the Skills and Interest MONitor or SIMON (Fonteyne, 2017), was specifically designed to guide or reorient high school graduates towards appropriate higher education study programs. The instrument attempts to predict academic achievement based on a battery of tests or questionnaires that scores students on a number of established (non-) cognitive predictors of academic achievement. Scoring below a certain threshold, students can then be dissuaded to enroll for a specific program. Indeed, a study by Fonteyne and colleagues (2018) regarding feedback on study choice found that negative feedback primarily lead to more study choice disengagement instead of continued engagement. However, literature and most practical applications almost entirely focus on modeling study success, and not on predicting study failure. For sure, by focusing exclusively on students that will almost certainly complete the program and using high thresholds, an important part of student potential is discarded without further consideration, even though a lot more students could have had a fair chance of completing the program to begin with. Moreover, educational and psychological literature almost exclusively focus on explaining population variance in

academic achievement, rather than predicting individual results (Shmueli, 2010). Indeed, predictors of academic achievement are generally used to explain the variance in study results and not to individually identify who will pass and who will not. The present study wants to add to literature on academic achievement by investigating how well we can identify students that are prone to study failure in a specific program, while still correctly identifying successful students. For this purpose, we will construct a model that specifically focuses on prediction of individual student outcomes. This way, only the students that lack fundamental skills or properties to complete the specific program are the target of reorientation or extra schooling advice, while the others are given a fair chance to enroll or to continue in a freely chosen study program.

Predictors of Academic Achievement

As literature already harbors a rich tradition of studies on the predictors of academic achievement, it is beyond the scope of the present study to provide a full overview. For a concise overview including (combined) effect sizes of the different predictors, we refer to the review of meta-analyses by Schneider and Preckel (2017). They found that successful students are characterized by high prior achievement (study antecedents) and intelligence, high self-efficacy, conscientiousness, and the goal-directed use of learning strategies. For the present study, we have thus used an operationalization of those variables, while also including variables like vocational interests and motivation as recent meta-analytic research has indicated that their importance towards study achievement is underestimated (see below).

A first major predictor of academic achievement is cognitive ability, which is considered as the strongest predictor of academic achievement, with reported correlations ranging from $r = .30$ to $r = .70$ (for an overview and meta-analytic evidence, see Roth et al., 2015 and Rohde & Thompson, 2007). Literature already displays a vast array of research on different iterations of the concept besides cognitive ability like intelligence and fluid g

(Gottfredson, 2002; Jensen, 1998; Vernon, 1983). Cognitive ability and its influence on study achievement also instigated the creation of instruments like SAT (historically called the Scholastic Aptitude or Assessment Test, see also <https://collegereadiness.collegeboard.org/>) and ACT (American College test, see also <http://www.act.org/>) that are used in the United States for college admission. In such a test, cognitive ability is operationalized as a mix of language (comprehensive reading and writing), mathematics (with and without calculator) and science. A study on the relationship between the *g* factor and the SAT scores revealed correlations up to $r = .86$, providing evidence that these tests are in fact measuring cognitive ability (Frey & Detterman, 2004). The cognitive tests within the SIMON project were designed to try and identify those students that were at risk of failing due to a lack of basic skills, making the overall difficulty rather low. For the present study, we have thus included data on a mathematics test, a comprehensive reading test and a vocabulary test. For some more typically scientific programs like bio-engineering, the data also contained an advanced mathematics test (instead of the base test), a chemistry test and a physics test. Pilot testing had revealed that the initial mathematics test was too easy (causing unwanted range restriction and ceiling effects) and that a topic specific test could add incremental validity towards predicting the individual results of students in addition to advanced mathematics.

A second major predictor of academic achievement in higher education are the student's study antecedents. Though a student's study antecedents are often strongly related to cognitive ability, study antecedents do have criterion validity in their own right. As an example, a large national study in New Zealand on the performance of students in medical school indicated that students' admission GPA (during their first year of college) explained up to 35% of the variance in academic achievement during the six year cycle (Poole, Shulruf, Rudland, & Wilkinson, 2012). In addition, a recent report from the Department of Education in the United States found that high school GPA explained higher education college grades

even better than standardized, cognitive based exams like the SAT or the ACT (Hodara & Lewis, 2017). Indeed, high school GPA explained 9 to 18% of the variance in college grades, while standardized exams only explained 1 to 5%. Moreover, a meta-analysis featuring data of over 200,000 students, revealed that at lower levels of higher education GPA, high school GPA was more explanative towards higher education GPA than the scores on the composite ACT (Noble & Sawyer, 2002). This is an important finding as the present study wants to specifically target students at risk of failure which inherently have lower levels of GPA. However, despite its obvious validity in explaining study results, GPA has been the subject of an ongoing debate whether such a measure is valid and reliable over different institutions or teachers (Graham, 2015; Harris, 2003; Sackett, Borneman, & Connelly, 2008). In response, several alternatives for GPA have been proposed. One such example is the mathematics level for students in high school. Indeed, in an open access environment study on achievement in STEM fields, the student's mathematics level at high school was related to first year higher education GPA, explaining about 4% of the variance in study results (Pinxten, Van Soom, Peeters, De Laet, & Langie, 2017). The same study also reports the complete study background in high school (including GPA) explained up to a quarter of the variance in higher education academic achievement. For the present study, we have thus opted to include the number of hours of mathematics in the senior year of high school as an index of mathematics level .

Although cognitive ability and a student's study antecedents can already explain a substantial proportion of the variance in academic achievement, other non-cognitive predictors also show incremental validity. As a third major predictor, motivation is a broad concept that harbors quite a few constructs. A meta – analysis on these constructs has shown that motivation can explain up to 4% of the variance in school results, even when controlling for cognitive ability (Kriegbaum, Becker, & Spinath, 2018). Within the scope of the present

study, the motivation concept is approached through the Self Determination Theory (Ryan & Deci, 2000; Ryan & Deci, 2017). According to this theory, two types of motivation (on a continuum from intrinsic to extrinsic) can direct the behavior of students: autonomous (driven by personal importance) and controlled (driven by external factors) motivation (Vansteenkiste, Sierens, Soenens, Luyckx, & Lens, 2009). As an example, a student can have good grades because he or she likes studying the chosen courses (autonomous motivation) or because he or she does not want to disappoint his or her parents (controlled motivation). Both autonomous and controlled motivation are included in our study.

A fourth major predictor of academic achievement is vocational interest. To this date, the RIASEC model is one of the most influential models to describe vocational interests in higher education (Holland, 1997; Nauta, 2010). At its core, the hexagonal model consists of six interest dimensions that are organized in a specific clockwise order: realistic, investigative, artistic, social, enterprising and conventional (Tracey & Rounds, 1995). Resemblance to these interest types is examined by a range of RIASEC instruments (Arbona, 2000; Spokane, Meir & Catalano, 2000; Toomey, Levinson, & Palmer, 2009). As an example, the SIMON-I questionnaire was specifically tailored to provide custom made orientation advice towards study programs for graduated high school students making the transition to higher education (Fonteyne, Wille, Duyck, & De Fruyt, 2017). Allen and Robbins (2010) also suggested a method to obtain specific program profiles by averaging RIASEC scores of all students in a specific program. Moreover, the RIASEC model has the intrinsic property of depicting interest profiles of both students and programs on the same hexagon as commensurate measures, so that comparisons between both become possible. Indeed, the RIASEC hexagon allows to establish how good the person - environment fit (PE fit) between a student and a study program actually is by comparing both profiles (Rounds & Su, 2014). This is a crucial feature because there is still debate in literature on how large the influence of

interest PE fit actually is. Indeed, correlation estimations for academic grades are ranging from close to zero to up to $r = .31$, also partly depending on the context of the study (Nye, Su, Rounds, & Drasgow, 2012, Schelfhout et al., 2019b). For the present study, the student RIASEC scores were obtained using the SIMON-I (Fonteyne, Wille, et al., 2017). Two continuous measures of PE fit were derived from these RIASEC scores: Euclidean distance and correlation fit. Euclidean distance measures the distance between student and program RIASEC profiles in a two-dimensional plane using the Prediger dimensions of People / Things and Data / Ideas (Prediger, 1982; Wille, Tracey, Feys & De Fruyt, 2014). In contrast, correlation fit does not focus on distance, but on the pattern of both profiles by calculating the Pearson's product moment correlation coefficient between the scores on the six dimensions of both the student profile and the program profile (Schelfhout et al., 2019a). Both continuous measures have explanative power towards academic achievement (Tracey, Allen, & Robbins, 2012) and can be used for program-specific environments.

Besides these four major predictors of academic achievement, there are a number of other constructs that have been (partly) incorporated into the present study due to their potential and availability, such as the conscientiousness dimension of the Big Five personality traits (De Fruyt & Mervielde, 1996; Furnham, Chamorro-Premuzic, & McDougall, 2003). According to Costa and McCrae (1992), conscientiousness incorporates six facets of personality: achievement-striving, competence, deliberation, dutifulness, order and self-discipline. Meta-analytic research indicates that high scores on this dimension are related to higher grades, with correlations reaching $r = .27$ or an equivalent explained variance of about 7% (Trapmann, Hell, Hirn, & Schuler, 2007). Also, the construct of self-control is related to conscientiousness, but it has a very distinct description through the aspect of conflict. Indeed, Duckworth, Taxer, Eskreis-Winkler, Galla and Gross (2019) describe self-control as the self-regulation of conflicting short and long term goals. Duckworth and colleagues (2019) also

discussed the explanative power towards academic persistence and performance, with correlations reaching $r = .19$, independent of cognitive ability. Finally, recent meta-analytic research on the correlates of academic performance investigated some other promising non-cognitive predictors of academic achievement including academic self-efficacy, (cognitive) test anxiety and metacognition (Richardson, Abraham, & Bond, 2012). Academic self-efficacy can be described as a belief in the personal capacity to organize and execute actions in order to attain desired goals (Bandura, 1993). Cognitive test anxiety is the anxiety students may experience in learning and evaluation contexts, and can lead to negative outcomes like less academic success (Credé & Kuncel, 2008). Metacognition can be described as self-knowledge of self-regulation and motivation, which allegedly can help towards getting better study results (Kitsantas et al., 2008). Richardson and colleagues (2012) thus reported small to medium sized correlations between university grades and academic self-efficacy, test anxiety and metacognition of $r = .28$, $r = -.21$ and $r = .14$ respectively.

Modeling Academic Achievement

In psychology and educational science, predictors of academic achievement are usually molded into some form of linear or logistic regression model to evaluate how well the (combined) predictors can explain the population variance in study outcomes like GPA. Both the total explained population variance of the model as well as the incremental contributions of the predictors (when controlling for the others) are commonly reported. For instance, Nye, Butt, Bradburn, and Prasad (2018) recently found that the combination of cognitive ability (e.g., ACT scores), the scores on a situational judgement test, biodata, and interest congruence (PE fit) could explain about 33% of the variance in overall college GPA, with interest congruence providing 3% of unique explained population variance above and beyond the other predictors. The vast majority of research in psychology and educational science has adopted this explanative approach.

However, Shmueli (2010) has correctly pointed out that models with high amounts of explained population variance do not necessarily predict individual student results. In other words, explaining high amounts of population variance in study results does not automatically validate individual student prediction of which students will pass and which students will fail. Indeed, both applications of statistical modeling have been heavily conflated, with the vast majority of psychological and educational research focusing on using the explanative population approach. However, for the present study, we are less interested in explaining the population variance in student achievement, but want to predict which individuals will fail (or pass) their program. To achieve this goal in the present study and add to literature, we are modeling academic achievement using the Akaike's Information Criterion (AIC) that focuses on minimizing prediction errors of full models instead of explaining variance through individual predictors (Burnham & Anderson, 2002).

Moreover, these explanative models in literature are almost always run across programs instead of using a program-specific approach. As an exception, Fonteyne, Duyck and De Fruyt (2017) did make use of a program-specific approach towards academic achievement as different programs can have different patterns of important predictors. Results showed that their program-specific approach had a clear edge over a more general approach, with each program featuring a unique combination of predictors towards academic achievement. As such, the study explained 23% of the variance in pass rates and GPA across programs, with a program-specific maximum of 29% for pass rates and 28% for GPA. Though the main focus of the study was on explaining population variance, the study also managed to correctly identify an average of 13% (with a maximum of 26.5%) of the students that were prone to failing their program (true positives), while tolerating only a small amount (5%) of false positives (i.e., students that passed who were identified as failing). In sum, the study provided much needed criterion validity for individual student prediction through the

combined use of study antecedents, cognitive ability, motivation, conscientiousness, test anxiety and self-efficacy (with also negative effects). The only predictor that did not contribute to explanation or prediction was metacognition. The study did have its limitations as the focus and methodology of the study was oriented towards explanation of population variance. Indeed, a good prediction of individual student results was considered the consequence of explaining a lot of population variance, not a goal on its own. For the present study, we want to specifically target individual prediction of student results by using a more appropriate focus and methodology. Also the dataset for the present study is nearly three times as large and as wide ($N = 6,624$ over 21 programs), increasing the external validity of the results.

Present Study

The present study wants to add to the literature on academic achievement by investigating how well we can identify students that are prone to failing in a specific open access program, while still correctly identifying passing students. As such, we allow for a strict 5% false positive rate (passing students identified as failing) in the following hypothesis,

H₁: *The program-specific AIC model identifies students at risk of failing above chance level.*

However, as we also want to explore the possibility of identifying even more failing students, we will investigate slightly more lenient boundaries at 10%, 15% and 20% of false positives. As such, we make the following hypothesis,

H₂: *The program-specific AIC model shows significant gains over programs in identifying students at risk of failing using more lenient false positive rates.*

Next, we want to provide an indication just how large the impact of the individual predictors is towards identifying students at-risk of failing. Moreover, as identifying students at risk of failing is a very specific binary operationalization of academic achievement, we have also

included GPA as a continuous operationalization of academic achievement. However, we cannot merely use explained variance as a measure for prediction as we have already discussed its conflated use (Shmueli, 2010). Still, as we are using a program-specific approach, we can indicate in how many programs a predictor is a part of the final predictive AIC model. A predictor that features in a large number of programs thus indicates a more broad impact range. As such, we try to reject the following null hypotheses,

H₃: *Study antecedents, cognitive ability, motivation, vocational interest, conscientiousness, self-control, academic self-efficacy, test anxiety and metacognition have an individual impact on the prediction of academic achievement over programs.*

For matters of comparison, we have also included the explained population variance for all program-specific models.

4.3. Method and Materials

The SIMON Project in an Open Access Study Environment

The present study was conducted within the context of the SIMON project that is currently still running (Fonteyne, 2017). Each prospective student with a high school degree can freely choose any program at university (with exception of medicine, dentistry and performing arts), without passing a standardized exam or meeting a certain level of high school GPA. All students have to pay a relatively modest yearly tuition fee of about € 920 or about \$ 1,050. Underprivileged students are also entitled to various government scholarships. In this open access higher education academic environment, the project aims to dispense program-specific orientation advice for each student prior to, and during the first year of an academic bachelor. For instance, if students have really low chances of success, they can be advised to upgrade their basic skills or reorient towards a more suitable program. SIMON's ultimate goal is to improve student study success. This is important because less than 40% of the first year students manages to fully pass their first year program and stay on the model

trajectory in a three-year bachelor program (Fonteyne, Duyck, & De Fruyt, 2017). The orientation advice is based on validated test results measuring the basic skills to succeed in a specific higher education program, using established predictors of academic achievement as discussed in the introduction. Only those students that lack these basic skills are advised to take extra schooling in order to upgrade their skill set or to reorient towards a different program altogether.

Data

The dataset for this study was extracted from the data available within the SIMON project. The present dataset thus includes SIMON test results (dating the start of the first academic year) and subsequent exam results (dating the end of the first semester and second semester) of students ($N = 6,624$) across 21 programs of a West-European open access university from 2016 to 2018. Table 1 shows an overview of these programs. We only incorporated the programs with $n > 100$ students to ensure the models had sufficient input to begin with, given the number of predictors. The RIASEC profiles for the 21 programs were obtained from an independent dataset ($N_0 = 6,572$) by averaging out the scores of successful and persistent students on all six dimensions as described by Allen and Robbins (2010).

Table 1. Study Programs.

number	program	students
1	psychology	834
2	communication science	189
3	educational science	204
4	political science	143
5	law	264
6	criminology	293
7	speech language and hearing science	101
8	linguistics	284
9	history	112
10	veterinary medicine	347
11	rehabilitation science and physiotherapy	527
12	pharmaceutical science	361
13	bioscience engineering	377

14	economical science	560
15	biomedical science	229
16	civil engineering	388
17	business administration	450
18	bioscience	118
19	industrial engineering	515
20	applied linguistics	219
21	biochemistry and biotechnology	109

Measures

Table 2 gives an overview of the pool of variables used to construct our predictive models. As the reliability and validity of a measure is dependent on both sample and instrument, we have checked the reliability and validity for all measures regarding the current sample (Harris, 2003; Graham, 2015). For a complete description of these measures, their reliability and their validation, we refer to APPENDIX A.

Table 2. Pool of Available Predictors.

Variables
dependent variables
<i>PASS</i>
<i>GPA</i>
study background
<i>high school mathematics package</i>
<i>high school GPA</i>
cognitive ability
<i>vocabulary</i>
<i>comprehensive reading</i>
<i>chemistry</i>
<i>physics</i>
<i>mathematics - baseline</i>
<i>mathematics - advanced</i>
motivation
<i>autonomous motivation</i>
<i>controlled motivation</i>
vocational interest
<i>realistic dimension</i>
<i>investigative dimension</i>
<i>artistic dimension</i>
<i>social dimension</i>
<i>enterprising dimension</i>

<i>conventional dimension</i>
<i>Euclidean distance</i>
<i>correlation fit</i>
<hr/>
<i>other predictors</i>
<hr/>
<i>conscientiousness</i>
<i>metacognition-knowledge</i>
<i>metacognition-regulation</i>
<i>cognitive test anxiety</i>
<i>self-control</i>
<i>academic self-efficacy comprehension</i>
<i>academic self-efficacy effort</i>
<hr/>

Note. GPA = grade point average. In order for a student to obtain a PASS (0/1) for his first year curriculum, the student has to individually pass all courses.

Procedure and Analyses

We have constructed predictive models for both PASS (**H1** to **H3**) and GPA (**H3**) using an AIC stepwise selection procedure to select the best predicting model (Burnham & Anderson, 2002). Tables 3 and 4 show the final program-specific models for predicting PASS and GPA. As we are dealing with relatively small sample sizes at the level of specific programs, it is crucial to avoid overfitting data noise as an actual effect towards prediction (Shmueli, 2010). As such, we have used the more strict version of the AIC including a correction (based on the sample size) for each predictor that enters the model (Cavanaugh, 1997). By comparing all possible predictor combinations (with linear and quadratic terms) against each other, the combination with the smallest chance of information loss is selected as the final model. Such a model limits the information loss by minimizing the prediction error for each student, which is the equivalent of a leave-one-out-cross-validation methodology. In such a methodology, a model is trained on all data minus the data for one specific student. Subsequently, the prediction of the model is tested against the actual outcome for each student separately. Finally, the model that renders the smallest prediction error across all students is selected. Although these models were specifically designed towards individual student prediction, the models can also be used to explain the population variance of study results in their programs. As the models were cross-validated against model overfitting, the explained

population variance will be rather conservative. To test how many failing students (true positives) we can identify at a certain rate of false positives (passing students falsely predicted to be failing) we need to balance both variables. A receiver operating characteristic curve or ROC curve balances the true positive rate (sensitivity) versus the false positive rate (1-true negative rate or 1-specificity). We have used a ROC curve to establish the sensitivity for all 21 specific program models at a false positive rate of 5, 10, 15 and 20%. As an example of a well-performing program-specific model, Figure 1 shows the ROC curve for biochemistry and biotechnology (program 21). As a measure of how well we can distinguish passing and failing students in each program, the area under the curve (AUC) is also reported in Table 3.

Table 3. Program-Specific AIC Models Predicting PASS.

program	model	R ²	AUC
1	$-2.57 + 0.002 \times \text{high school GPA}^2 + 0.01 \times \text{mathematics}^2 + 1.75 \times \text{high school mathematics package} - 0.16 \times \text{high school mathematics package}^2 - 0.18 \times \text{high school GPA} - 0.02 \times R - 0.004 \times \text{cognitive test anxiety}^2$.26	.75
2	$33.41 + 0.01 \times \text{high school GPA}^2 + 3.03 \times \text{high school mathematics package} - 0.0001 \times E^2 - 1.27 \times \text{high school GPA} - 0.27 \times \text{high school mathematics package}^2$.30	.74
3	$-20.5 + 0.15 \times \text{high school GPA} + 0.044 \times S - 0.00032 \times E^2 + 3.15 \times \text{high school mathematics package} - 0.3 \times \text{high school mathematics package}^2$.29	.74
4	$-11.76 + 0.00012 \times \text{high school GPA}^2 + 0.009 \times \text{autonomous motivation}^2 + 0.52 \times \text{high school mathematics package} + 0.024 \times A$.41	.81
5	$-22.55 + 0.0021 \times \text{high school GPA}^2 + 3.38 \times \text{high school mathematics package} + 0.2 \times \text{autonomous motivation} - 0.29 \times \text{high school mathematics package}^2$.48	.85
6	$-9.61 + 0.00078 \times \text{high school GPA}^2 + 0.000088 \times \text{conscientiousness}^2 + 0.16 \times \text{mathematics} + 0.0052 \times \text{controlled motivation}^2$.23	.72
7	$1.69 - 0.012 \times \text{cognitive test anxiety}^2$.10	.65
8	$-9.91 + 0.0014 \times \text{high school GPA}^2 + 0.15 \times \text{autonomous motivation}$.36	.80
9	$-0.45 + 0.0018 \times \text{high school GPA}^2 + 0.018 \times \text{vocabulary}^2 - 0.38 \times \text{mathematics} - 0.34 \times \text{cognitive test anxiety} + 0.6 \times \text{academic self-efficacy comprehension} + 0.67 \times \text{high school mathematics package}$.53	.85
10	$-5.67 + 0.00061 \times \text{high school GPA}^2 + 0.12 \times \text{physics} + 0.3 \times \text{high school mathematics package} - 0.12 \times \text{cognitive test anxiety}$.26	.75
11	$-8.56 + 0.0011 \times \text{high school GPA}^2 + 0.17 \times \text{physics}$.28	.77
12	$-10.91 + 0.00085 \times \text{high school GPA}^2 + 0.28 \times \text{mathematics} + 0.11 \times \text{physics}$.28	.75
13	$3.56 + 0.0032 \times \text{high school GPA}^2 + 0.0068 \times \text{mathematics}^2 - 0.31 \times \text{high school GPA}$.31	.76
14	$-16.35 + 0.2 \times \text{high school GPA} + 0.18 \times \text{mathematics}$.36	.80
15	$-15.41 + 0.14 \times \text{high school GPA} + 0.24 \times \text{self-control} + 0.12 \times \text{autonomous motivation} + 0.0045 \times \text{physics}^2$.31	.76
16	$-12.55 + 0.13 \times \text{high school GPA} + 0.17 \times \text{mathematics}$.25	.74
17	$-13.38 + 0.15 \times \text{high school GPA} + 0.14 \times \text{mathematics} + 0.34 \times \text{high school mathematics package}$.25	.75
18	$-14.32 + 0.13 \times \text{high school GPA} + 0.54 \times \text{high school mathematics package} + 0.01 \times \text{self-control}$.30	.75
19	$0.63 + 0.00015 \times \text{high school GPA}^2 + 0.021 \times \text{mathematics}$.13	.70
20	$-9.43 + 0.0011 \times \text{high school GPA}^2 + 0.0059 \times \text{mathematics}^2 + 0.094 \times \text{comprehensive reading}$.31	.77
21	$-8.74 + 0.0023 \times \text{high school GPA}^2 + 0.33 \times \text{mathematics} - 0.51 \times \text{chemistry} + 0.014 \times \text{physics}^2 - 2.62 \times \text{correlation fit}^2$.68	.91

Note. A student obtaining a PASS for his first year curriculum, has to individually pass all courses. The model is a logistic model, of which only the linear prediction element is shown. The explained population variance was measured through a Nagelkerke's (pseudo) R^2 . The distinctive power of the model identifying failing from passing student is indicated by the Area Under the Curve (AUC). GPA = grade point average, R = realistic interest dimension, I = investigative interest dimension, A = artistic interest dimension, S = social interest dimension, E= enterprising interest dimension, C = conventional interest dimension. The order of the terms is displayed as originally rendered by the algorithm and represents the relative importance of the terms towards prediction. The terms can reflect linear effects, pure quadratic effects or curvilinear effects (linear + quadratic

effect). It is also possible that a term acts as a suppressor effect. For instance, *correlation fit*² in its own right has a positive effect on the passing rate, but acts as a suppressor for other effects in the model (negative effect). However, an individual breakdown of each model is beyond the scope of this study.

Table 4. Program-specific models predicting GPA.

program	model	R ²	GPA error
1	$455.11 + 0.17 \times \text{high school GPA}^2 + 0.47 \times \text{mathematics}^2 + 86.59 \times \text{high school mathematics package} - 2.23 \times R - 14.33 \times \text{high school GPA} - 6.62 \times \text{high school mathematics package}^2 - 0.43 \times \text{cognitive test anxiety}^2 - 0.49 \times \text{academic self-efficacy comprehension}^2 + 3.40 \times \text{comprehensive reading} + 0.01 \times S^2$.33	109.20
2	$-417.17 + 0.06 \times \text{high school GPA}^2 + 6.63 \times \text{comprehensive reading} + 219.18 \times \text{high school mathematics package} - 20.54 \times \text{high school mathematics package}^2$.25	110.53
3	$-507.1 + 7.96 \times \text{high school GPA} + 144.3 \times \text{high school mathematics package} - 13.2 \times \text{high school mathematics package}^2 + 1.88 \times S + 4.35 \times \text{comprehensive reading} - 0.014 \times C2$.42	65.93
4	$-559.10 \times \text{high school GPA}^2 + 0.069 \times \text{mathematics}^2 + 0.51 \times \text{autonomous motivation} + 20.37 \times \text{high school mathematics package} + 28.30 \times \text{vocabulary}^2 - 0.041 \times I2 + 0.0031 \times \text{Euclidean distance}^2 - 0.0024 \times C^2$.43	119.48
5	$-780.3 + 0.11 \times \text{high school GPA}^2 + 121.9 \times \text{high school mathematics package} + 0.23 \times \text{mathematics}^2 - 0.038 \times R2 - 9.64 \times \text{high school mathematics package}^2 + 0.41 \times \text{vocabulary}^2 + 6.29 \times \text{autonomous motivation} + 17.34 \times \text{academic self-efficacy effort} - 12.34 \times \text{metacognition-knowledge}$.49	96.67
6	$72.75 + 0.050 \times \text{high school GPA}^2 + 10.10 \times \text{mathematics} + 0.011 \times S^2 + 0.0049 \times \text{conscientiousness}^2 + 0.46 \times \text{controlled motivation}^2 + 18.53 \times \text{high school mathematics package} - 11.26 \times \text{cognitive test anxiety} - 0.58 \times \text{academic self-efficacy comprehension}^2$.31	104.84
7	$599.77 + 11.53 \times \text{physics} - 0.74 \times \text{cognitive test anxiety}^2$.19	103.76
8	$-605.4 + 0.09 \times \text{high school GPA}^2 + 7.41 \times \text{autonomous motivation} + 12.86 \times \text{academic self-efficacy effort} + 0.19 \times \text{comprehensive reading} + 124.2 \times \text{high school mathematics package} - 11.69 \times \text{high school mathematics package}^2$.47	98.16
9	$382.05 + 0.10 \times \text{high school GPA}^2 + 0.98 \times \text{vocabulary}^2 - 37.64 \times \text{academic self-efficacy comprehension} + -23.54 \times \text{cognitive test anxiety} + 103.11 \times \text{correlation fit}^2$.50	111.40
10	$-328.8 + 497.1 \times \text{high school GPA}^2 + 11.06 \times \text{physics} + 96.27 \times \text{high school mathematics package} + 0.046 \times \text{conscientiousness}^2 - 6.96 \times \text{high school mathematics package}^2$.29	126.40
11	$-777.8 + 11.47 \times \text{high school GPA} + 7.47 \times \text{physics} + 182.1 \times \text{high school mathematics package} - 15.31 \times \text{high school mathematics package}^2 - 15.83 \times \text{academic self-efficacy comprehension} + 0.004 \times \text{conscientiousness}^2$.36	107.86
12	$-67.08 + 0.064 \times \text{high school GPA}^2 + 13.16 \times \text{physics} + 11.33 \times \text{chemistry} - 1.34 \times I$.32	114.13
13	$626.87 + 0.20 \times \text{high school GPA}^2 + 0.45 \times \text{mathematics}^2 - 17.64 \times \text{high school GPA}$.33	108.66
14	$-907.32 + 13.36 \times \text{high school GPA} + 8.94 \times \text{mathematics} + 139.14 \times \text{high school mathematics package} - 10.86 \times \text{high school mathematics package}^2 - 0.35 \times \text{cognitive test anxiety}^2$.42	98.04
15	$-263.6 + 9.62 \times \text{high school GPA} + 8.87 \times \text{mathematics} + 10.04 \times \text{physics} + 0.77 \times \text{academic self-efficacy effort} - 21.64 \times \text{academic self-efficacy comprehension}$.38	114.26
16	$-183.4 + 0.059 \times \text{high school GPA}^2 + 17.9 \times \text{mathematics} + 0.45 \times \text{academic self-efficacy comprehension}^2$.29	113.01
17	$-650.54 + 11.99 \times \text{high school GPA} + 8.33 \times \text{mathematics} + 19.02 \times \text{high school mathematics package} + 12.39 \times \text{vocabulary}$.28	103.81
18	$-541.65 + 10.61 \times \text{high school GPA} + 26.98 \times \text{high school mathematics package} - 0.06 \times I2 + 5.55 \times I + 9.18 \times \text{mathematics}$.42	86.68
19	$-651.3 + 0.053 \times \text{high school GPA}^2 + 9.38 \times \text{mathematics} + 166.1 \times \text{high school mathematics package} + 12.51 \times \text{academic self-efficacy effort} - 11.54 \times \text{high school mathematics package}^2$.24	119.20
20	$-477.02 + 0.1 \times \text{high school GPA}^2 + 194.67 \times \text{high school mathematics package} - 18.49 \times \text{high school mathematics package}^2$.36	110.16
21	$-1906 + 0.1 \times \text{high school GPA}^2 + 8.43 \times \text{mathematics} - 30.23 \times \text{high school mathematics package} + 9.58 \times \text{physics} + 190 \times \text{academic self-efficacy effort} - 5.95 \times \text{academic self-efficacy effort}$.67	86.72

Note. The model is a linear model (in its parameters). The explained population variance was measured through an R². The accurateness of the model is indicated through a grade point average (GPA) error, the average absolute error in predicting study achievement for that specific program. The normal average error over students for each program is not significantly different from zero, indicating that our estimator of GPA error is in fact pure. GPA = grade point average, R = realistic interest dimension, I = investigative interest dimension, A = artistic interest dimension, S = social interest dimension, E= enterprising interest dimension, C = conventional interest dimension. The order of the terms is again displayed as originally rendered by the algorithm and represents the relative importance of the terms towards prediction. The terms can reflect linear effects, pure quadratic effects or curvilinear effects (linear + quadratic effect). It is also possible that a term acts as a suppressor effect.

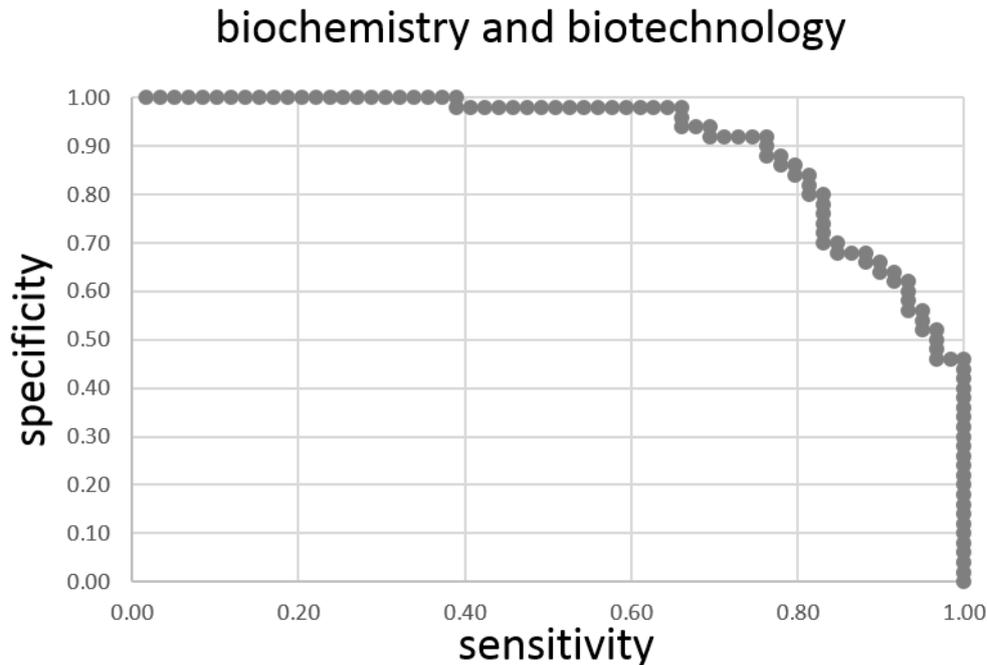


Figure 1. ROC Curve for the Program *biochemistry and biotechnology*. Sensitivity indicates the proportion true positives, or students correctly identified as failing. Specificity indicated the proportion true negatives, or students correctly identified as passing. The Area Under the Curve (AUC) indicates how well the model can distinguish between passing and failing students. For this model, the AUC amounts to 91%.

For **H1**, we have used a one-sample two-sided t-test. For **H2**, we have used a two-sample, two-sided t-test. For **H3**, predictors that are featuring in one or more specific program models are considered as predictive for the results of individual students. To obtain a more continuous measure, we are also reporting the frequency with which the predictors occur across all 21 program models as an indication of their range of impact. Effect sizes are reported using a Cohen's d ranging from a very small effect of $d = 0.01$ to a huge effect of $d = 2.00$ (Sawilowsky, 2009). Finally, we have calculated the correlation between the average errors in GPA over programs (as an indication of individual student prediction) and the explained population variance in each program to have an indication how individual student prediction and explained population variance relate to each other for the current data sample.

4.4. Results

Table 5 shows the percentage of identified failing students for each program at a 5%, 10%, 15% and 20% false positive rate as determined through the program-specific AIC models in Table 3. For H1, a one-sample, two-sided Welch t-test revealed a significant difference from zero, $t(20) = 10.98, p < .001$, with a huge effect size of $d = 2.40$. As predicted, we have indeed found that the program-specific AIC model successfully identifies about 29% of all students at risk of failing at a 5% false positive rate.

Table 5. Identified Students at Risk of Failing across all Programs.

program	5% FPR	10% FPR	15% FPR	20% FPR
1	27	40	49	57
2	17	25	31	57
3	22	32	46	51
4	45	53	65	66
5	43	54	68	74
6	21	26	44	47
7	9	20	27	41
8	40	51	59	64
9	31	61	71	71
10	25	36	52	57
11	31	50	57	61
12	27	38	48	52
13	28	33	39	51
14	29	44	57	63
15	36	48	59	60
16	21	42	47	57
17	26	41	47	53
18	16	42	49	56
19	25	32	41	41
20	26	44	52	56
21	66	76	78	81
average	29	42	52	58

Note. FPR = false positive rate or the percentage of passing students identified as failing.

For H2, we tested the increase in identifying students at risk by applying more lenient false positive rates. Figure 2 shows the increase from a 5% to a 20% false positive rate. A two-sample, two-sided dependent Welch t-test revealed that there is a large (from 15% to 20%) to huge increase (from 5% to 10% and from 10% to 15%) at each rate. The biggest

absolute increase in identifying students at risk of about 13% occurs between the 5% and 10% false positive rate. Also, the biggest relative effect occurs between the 10% and 15% false positive rate, though the variance over programs regarding the gain in identifying students at risk is much larger at the 15% rate. Between the 15 and 20% false positive rate, both the absolute increase as the effect of identifying students at risk are less pronounced.

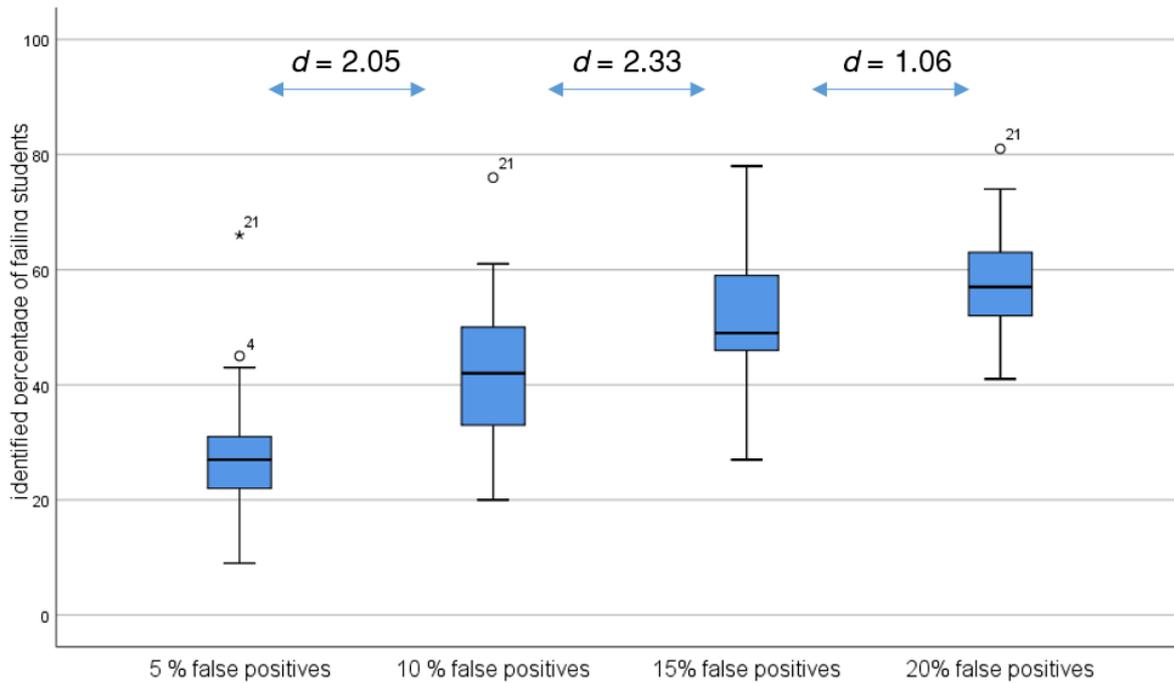


Figure 2. Boxplot of Identified Failing Students. False positive rates (FPR) at 5%, 10%, 15% and 20% indicating how many passing students are identified as failing. All effect sizes were significant, $p < .001$. Program 21 had the best performing model, making it an outlier at the 5%, 10% and 20% false positive rate.

Tables 6 and 7 show the frequency with which the different predictors occur in program-specific models for PASS and GPA respectively. For H3, our hypothesis was largely confirmed. Study antecedents, cognitive ability, motivation, vocational interest, conscientiousness, self-control, academic self-efficacy, test anxiety and metacognition predicted academic achievement for both the very specific PASS measure as the more general GPA measure to various degrees. For self-control and metacognition, we observe mixed evidence. Self-control displays predictive value for PASS but not for GPA, while we observe

the reverse pattern for metacognition. For the impact range of predictors, the correlation between predictor frequencies for GPA and PASS correlated $r = 0.89, p = .001$.

Table 6. Impact of PASS-Predictors Across Programs.

program	study antecedents	cognitive ability	motivation	vocational interest	conscientiousness	metacognition	cognitive test anxiety	self-control	academic self-efficacy
1	1	1	0	1	0	0	1	0	0
2	1	0	0	1	0	0	0	0	0
3	1	0	0	1	0	0	0	0	0
4	1	0	1	1	0	0	0	0	0
5	1	0	1	0	0	0	0	0	0
6	1	1	1	0	1	0	0	0	0
7	0	0	0	0	0	0	1	0	0
8	1	0	1	0	0	0	0	0	0
9	1	1	0	0	0	0	1	0	1
10	1	1	0	0	0	0	1	0	0
11	1	1	0	0	0	0	0	0	0
12	1	1	0	0	0	0	0	0	0
13	1	1	0	0	0	0	0	0	0
14	1	1	0	0	0	0	0	0	0
15	1	1	1	0	0	0	0	1	0
16	1	1	0	0	0	0	0	0	0
17	1	1	0	0	0	0	0	0	0
18	1	0	0	0	0	0	0	1	0
19	1	1	0	0	0	0	0	0	0
20	1	1	0	0	0	0	0	0	0
21	1	1	0	1	0	0	0	0	0
frequency	0.95	0.67	0.24	0.24	0.05	0.00	0.19	0.10	0.05

Note. Predictors are marked 1 if they occur in the predictive model of the program. The *frequency* indicates how large the impact range is of the predictor over programs.

Table 7. Impact of GPA-Predictors Across Programs.

program	study antecedents	cognitive ability	motivation	vocational interest	conscientiousness	metacognition	cognitive test anxiety	self-control	academic self-efficacy
1	1	1	0	1	0	0	1	0	1
2	1	1	0	0	0	0	0	0	0
3	1	1	0	1	0	0	0	0	0
4	1	1	1	1	0	0	0	0	0
5	1	1	1	1	0	1	0	0	1
6	1	1	1	1	1	0	1	0	1
7	0	1	0	0	0	0	1	0	0
8	1	1	1	0	0	0	0	0	1
9	1	1	0	1	0	0	1	0	1
10	1	1	0	0	1	0	0	0	0
11	1	1	0	0	1	0	0	0	1

12	1	1	1	0	0	0	0	0	0
13	1	1	0	0	0	0	0	0	0
14	1	1	0	0	0	0	1	0	0
15	1	1	0	0	0	0	0	0	1
16	1	1	0	0	0	0	0	0	1
17	1	1	0	0	0	0	0	0	0
18	1	1	0	1	0	0	0	0	0
19	1	1	0	0	0	0	0	0	1
20	1	0	0	0	0	0	0	0	0
21	1	1	0	0	0	0	0	0	1
frequency	0.95	0.95	0.24	0.33	0.14	0.05	0.24	0.00	0.48

Note. Predictors are marked 1 if they occur in the predictive model of the program. The *frequency* indicates how large the impact range is of the predictor over programs.

Finally, the correlation over programs (GPA) between explained population variance and average student prediction error amounted to $r = -.48$, $p = .03$. Though the correlation indicates a somewhat strong negative relation, GPA population variance over programs only explains about 23% in the GPA prediction error of individual students over programs.

4.5. Discussion

In an open access environment, everyone with a high school degree can enter almost any program. Early knowledge about lacking skills or non-fitting properties could prevent study failure and would improve the present low program success rates that many Western-European higher education systems show (Fonteyne, Duyck, et al., 2017; Fonteyne, 2017). By identifying these students at risk of failing at the beginning of their model trajectory, an advice towards extra schooling or towards reorientation can instigate student action, either by pursuing skill development or by reorienting to a more appropriate program altogether. In order to facilitate this advice, the present study set forward to investigate how well we can identify students prone to study failure across all programs of a big university, while still correctly identifying passing students. For this purpose, a program and prediction specific AIC model was applied to data at the start of the academic year and linked to outcomes of passing and GPA.

We have found that using a mainstream and easy to administer set of predictors (Schneider & Preckel, 2017), such a model can identify about 29% of the students at risk of failing their exams at a very strict false positive rate (5%), with program-specific peaks of up to 66%. At the most lenient false positive rate (20%), the percentage of identified students at risk shows a huge increase of up to on average 58%, with program-specific peaks of up to 81%. We also observe that the highest gains in identification occur between the 5% and 10% false positive mark (largest absolute increase of 13%) and the 10% and 15% false positive mark (largest effect size). However, the gains between the 10% and 15% mark have a much larger variance. We also found that on average, our models could explain up to 37% of the population variance (deviance) in academic achievement. These results thus make a number of important contributions to literature and practice.

First and foremost, to our knowledge, this was the first study in literature with the specific aim of identifying students at risk of failing, by using a methodology with a program-specific focus on predicting individual student results. Indeed, our program-specific approach seems duly warranted, as we found that the explained population variance in study results across programs only explained about 23% in the individual student error rate on the prediction of GPA. As Shmueli (2010) suggested, explanation and prediction are in fact related, but far from identical applications of statistical modeling. For the present data, this means that program models that show similar explained variance may display different individual prediction accuracy and vice versa. An assessment on the appropriate use regarding both applications of statistical modeling should therefore at least be considered for all future studies on academic achievement.

Second, the present study has established criterion validity on predicting individual student outcomes for a wide range of predictors through their impact range. This is somewhat different compared to the common approach in literature that focuses on explained population

variance. Indeed, literature already provided ample evidence towards the explained variance effects on academic achievement of study antecedents, cognitive ability, motivation, vocational interest, conscientiousness, self-control, academic self-efficacy, cognitive text anxiety and metacognition (Schneider & Preckel, 2017). Also, some studies found that these predictors not only provide an explanation of population variance, but also that they heavily influence the individual prediction of student results (Fonteyne, 2017; Fonteyne, Duyck, et al., 2017). Unfortunately, the predictive power was established while still using an explanative focus and methodology. For the present study, we tried to add to the literature by using a program-specific predictive model using linear and curvilinear terms in an attempt to replicate these findings. We included both a very specific measure (PASS rate of students) and a more general measure (GPA) of academic achievement. Our study thus found that both measures rendered largely the same results in impact range as indicated by the large correlation between both ($r = .89$). Study antecedents, cognitive ability, motivation, vocational interest, conscientiousness, self-control, academic self-efficacy, cognitive text anxiety and metacognition were all part of the program-specific models to various degrees. Study antecedents and cognitive ability had an impact on the largest range of programs (up to 95% of all programs), though cognitive ability was slightly less present regarding prediction using the more specific PASS measure. These findings on impact range are largely in line with reported results in literature on the explanative use of study antecedents (Noble & Sawyer, 2002; Poole et al, 2012). Moreover, the larger predictive impact range of study antecedents over cognitive ability for passing is analogous to findings on explanative power at lower levels of GPA (Hodara & Lewis, 2017). Vocational interest, motivation and test anxiety had a smaller impact range as they featured in about 20% to 30% of all program prediction models. These results on predictive impact range across programs are also largely in line with literature on the explanative use of predictors in showing incremental validity on top of study

antecedents and cognitive ability (Kriegbaum, 2018; Nye et al., 2018; Richardson et al., 2012; Credé & Kuncel, 2008; Trapmann et al., 2007). The analyses for both measures also revealed some minor differences. Self-control showed a small impact range on academic achievement using the specific PASS measure, but no impact at all for the more general GPA measure. In contrast, metacognition showed the opposite pattern. These findings are somewhat in line with results on the explanative power of these constructs above and beyond cognitive ability (Duckworth et al., 2019; Kitsantas et al., 2008). Finally, academic self-efficacy had a medium to large impact on GPA, while only having a very small impact over programs on GPA. Some of the program-specific effects were actually negative, which is in line with the findings of Fonteyne and colleagues (2017), but in contrast to literature (Bandura, 1993; Richardson et al., 2012). It seems that students in an open access study environment are sometimes overconfident regarding their (basic) skills.

Third, the results from our study are quite an upgrade compared to similar studies in an open access environment for both explained population variance and individual student prediction. As an example, compared to Fonteyne and colleagues (2017), our study explains about two thirds more variance (23% vs 37%), while identifying more than double the amount of students at risk of failing (13% vs 29%). We suspect that the increase in explained population variance is due to the inclusion of high school GPA as a form of prior study results, as we already know it is highly explanative towards academic achievement in higher education (Noble & Sawyer, 2002; Pinxten et al., 2017; Poole et al., 2012). Moreover, we suspect that the huge difference in individual student prediction is not only due to the inclusion of high school GPA, but also to the use of our prediction specific methodology as the difference towards prediction was much larger than the difference towards explanation.

Finally, our study also has important consequences for study counseling, especially through the use of more lenient false positive rates. Indeed, the analyses on more lenient false

positive rates could be used by educational policy makers to fine tune test score criteria to acceptable false positive rates, given desired true positives goals. Of course, our specific advice for the use of these different rates is primarily dependent on the outcome of sample specific analyses. Using our present study as an example, we could agree to a false positive rate of 10% as it provides a 45% relative increase in identifying failing students (from 29% to 42%). We would advise caution when using the 15% rate, as it causes more variance in the gains. We would not agree to an even more lenient false positive rate of 20% as this rate shows clear diminishing returns in identifying failing students. In contrast, counselors at college or university could also use a program-specific approach. As such, we recommend keeping the false positive rate at 5% or 10% at maximum for biochemistry and biotechnology (program 21) as more lenient rates would only increase the absolute gains in identifying students at risk by a mere 5%. However, we also recommend increasing the false positive rate for speech language and hearing science to even 20% as it provides an absolute gain of 32% and a relative gain of 356% in identifying students at risk.

As a final remark, as our study was conducted exclusively in an open access academic environment, future research also needs to further generalize our findings towards different, more constraint forms of higher education environments.

Conclusion

With the appropriate predictive and program-specific methodology, the present study has shown we can identify a large proportion of students that are prone to failing. Our study has also validated the use of predictors like study antecedents and cognitive ability for predicting individual student results alongside the already established explanation of population variance in academic achievement. In practice, study counselors can use this knowledge to advise students at risk to upgrade their basic skills or to reorient towards a different, more suitable program.

4.6. References

- Allen, J. & Robbins, S. (2010). Effects of Interest–Study program Congruence, Motivation, and Academic Performance on Timely Degree Attainment. *Journal of Counseling Psychology, 57*(1), 23-35. doi:10.1037/a0017267
- Arbona, C. (2000). Practice and research in career counseling and development – 1999. *Career Development Quarterly, 49*(2), 98-134. doi:10.1002/j.2161-0045.2000.tb00554.x
- Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist 28*(2), 117-148. doi:10.1207/s15326985ep2802_3
- Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A practical information-theoretic approach* (2nd ed.), Springer-Verlag.
- Cassady, J., & Finch, W. (2015). Using factor mixture modeling to identify dimensions of cognitive test anxiety. *Learning and Individual Differences, 41*, 14–20. <http://dx.doi.org/10.1016/j.lindif.2015.06.002>
- Cavanaugh, J. E. (1997). Unifying the derivations of the Akaike and corrected Akaike information criteria, *Statistics & Probability Letters, 31*, 201–208. doi:10.1016/s0167-7152(96)00128-9.
- Costa, P.T., & McCrae, R.R. (1992). *Revised NEO Personality Inventory and NEO Five Factor Inventory Professional manual*. Odessa; FL: Psychological Assessment Resources.
- 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

- De Fruyt, F., & Mervielde, Y. (1996). *Personality and interests as predictors of educational streaming and achievement*. *European Journal of Personality*, *10*(5), 405-425. doi: 10.1002/(SICI)1099-0984(199612)10:5<405::AID-PER255>3.0.CO;2-M
- De Fruyt, F., & Rolland, J. P. (2010). *PfPI. Beschrijving persoonlijkheid op het werk. Handleiding*. Amsterdam/Antwerpen: Pearson.
- Duckworth, A. L., Taxer, J. L., Eskreis-Winkler, L., Galla, B. M., & Gross, J. J. (2019). Self-control and Academic Achievement. *Annual Review of Psychology*, *70*, 373-393. doi: 10.1146/annurev-psych-010418-103230
- Fonteyne, L. (2017). *Constructing SIMON: a tool for evaluating personal interests and capacities to choose a post-secondary major that maximally suits the potential* (Doctoral dissertation). Retrieved 22nd of March 2019 from <https://biblio.ugent.be/publication/8535491>
- 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. doi: 10.1016/j.lindif.2017.05.003
- Fonteyne, L., Eelbode, A., Lanszweert, I., Roels, E., Schelfhout, S., Duyck, W., & De Fruyt, F. (2018). Career goal engagement following negative feedback : influence of expectancy-value and perceived feedback accuracy. *International Journal For Educational and Vocational Guidance*, *18*(2), 165–180.
- Fonteyne, L., Wille, B., Duyck, W., & De Fruyt, F. (2017). 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*, *17*(2), 233-262. doi:10.1007/s10775-016-9327-9

- Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science, 15*(6), 373-378. doi: 10.1111/j.0956-7976.2004.00687.x
- Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences, 14*(1), 49-66. doi:10.1016/j.lindif.2003.08.002
- Gottfredson, L. S. (2002). g: Highly general and highly practical. In R. J. Sternberg, & E. L. Grigorenko (Eds.), *The general factor of intelligence: How general is it?* (pp. 331–380). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Graham, S. (2015). Inaugural Editorial for the Journal of Educational Psychology. *Journal of Educational Psychology, 107*(1), 1-2. doi: 10.1037/edu0000007
- Harris, K.R. (2003). Editorial: Is the work as good as it could be? *Journal of Educational Psychology, 95*, 451-452. doi:10.1037/0022-0663.95.3.451
- Hodara, M., & Lewis, K. (2017). *How well does high school grade point average predict college performance by student urbanicity and timing of college entry?* (REL 2017–250). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest. Retrieved 27th of March 2019 from <http://ies.ed.gov/ncee/edlabs>.
- Holland, J.L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, FL: Psychological Assessment Resources.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- 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.

- Kriegbaum, Becker, & Spinath (2018). The relative importance of intelligence and motivation as predictors of school achievement: A meta-analysis. *Educational Research Review*, 25, 120-148. doi: 10.1016/j.edurev.2018.10.001
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44, 325–343. <http://dx.doi.org/10.3758/s13428-011-0146-0>
- Noble, J., & Sawyer, R. (2002). Predicting Different Levels of Academic Success in College Using High School GPA and ACT Composite Score. Retrieved 27th of March 2019 from www.researchgate.net/publication/237103886_Predicting_Different_Levels_of_Academic_Success_in_College_Using_High_School_GPA_and_ACT_Composite_Score
- Nye, C. D., Butt, S. M., Bradburn, J., & Prasad, J. (2018). Interests as predictors of performance: An omitted and underappreciated variable. *Journal of Vocational Behavior*, 108, 178-189. doi: 10.1016/j.jvb.2018.08.003
- 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
- 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.*
- Pinxten, M., Van Soom, C., Peeters, C., De Laet, T., & Langie, G. (2017). At-risk at the gate: prediction of study success of first-year science and engineering students in an open-admission university in Flanders—any incremental validity of study strategies? *European Journal of Psychology of Education*, 34(1), 45-66. doi: 10.1007/s10212-017-0361-x

- Poole P., Shulruf B., Rudland, J., & Wilkinson T. (2012). Comparison of UMAT scores and GPA in prediction of performance in medical school: a national study. *Medical Education* 46(2), 163-171. doi: 10.1111/j.1365-2923.2011.04078.x
- 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
- 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
- Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, 35, 83-92. doi:10.1016/j.intell.2006.05.004
- Roth, B., Becker, N., Romeyke, S., Schäfer, 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
- Rounds, J., & Su, R. (2014). The Nature and power of Interests. *Current Directions in Psychological Science*, 23(2), 98-103. doi: 10.1177/0963721414522812
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78.
doi: 10.1037/0003-066 x.55.1.68
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. New York: Guilford Publishing.
- Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). *High stakes testing in higher education and employment: Appraising the evidence for validity and fairness*. *American Psychologist*, 63(4), 215–227. doi: 10.1037/0003-066X.63.4.215

- Sawilowsky, S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, 8(2), 467-474.
- Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., De Fruyt, F., & Duyck, W. (2019a). From Interest Assessment to Study Orientation: an Empirical Advice Set Engine. *The Journal of Experimental Education*. doi: 10.1080/00220973.2019.1605327
- Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., De Fruyt, F., & Duyck, W. (2019b). The effects of vocational interest on study results: Student person – environment fit and program interest diversity, *PLoS ONE*. doi: 10.1371/journal.pone.0214618
- Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, 143(6), 565-600. doi: 10.1037/bul0000098
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19(4), 460–475. <http://dx.doi.org/10.1006/ceps.1994.1033>
- Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, 25(3), 289-310. doi: 10.1214/10-STS330
- Spokane, A.R., Meir, E.L., & Catalano, M. (2000). Person-environment congruence and Holland's theory: a review and reconsideration. *Journal of Vocational Behavior*, 57(2), 137-187. doi: 10.1006/jvbe.2000.177
- Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). "High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72(2), 271–324.
- Toomey, K.D., Levinson, E.M., & Palmer, E.J. (2009). A test of Holland's theory of vocational personalities and work environments. *Journal of Employment Counseling*, 46(2), 82-93. doi:10.1002/j.2161-1920.2009.tb00070.x2

- Tracey T. J. G., Rounds, J. (1997). 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. (1995). The Arbitrary Nature of Holland's RIASEC Types: a Concentric-Circles Structure. *Journal of Counseling Psychology*, 42(4), 431-439. doi: 10.1037/0022-0167.42.4.431
- 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
- Trapmann, S., Hell, B., Hirn, J.O.W., & Schuler, H. (2007). Meta-Analysis of the Relationship Between the Big Five and Academic Success at University. *Zeitschrift für Psychologie* 215(2), 132-151. doi:10.1027/0044-3409.215.2.132
- 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
- Vernon, P. A. (1983). Recent Findings on the Nature of g. Vernon, P. A. (1983). *Journal of Special Education*, 17(4), 389–400. doi:10.1177/002246698301700403
- 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

4.7. APPENDIX A:

Measures: Description, Reliability and Validation

To start, we have included two dependent measures of study success, obtained from exam results. *PASS* is a binary variable (0/1) and indicates if a student has passed all programs of his model trajectory and can fully advance to the next year. For the current data sample, 43% of the students fully passed their first year. *GPA* gives a more gradual indication of a student's academic achievement, presenting an average score on all courses ranging from 0 to 1000 ($M = 524.13$; $SD = 169.54$). Also, for a full discussion on the reliability and validity of GPA and PASS as dependent measures in a similar open access environment, we refer to Schelfhout and colleagues (2019b). However, for the present study, possible bias of these measures due to program-specific or even teacher-specific circumstances is eliminated altogether, as each program is modeled separately and all students have identical curricula within one program.

For study antecedents, we have included self-reported *high school GPA* ranging from 0 to 100 ($M = 72.08$; $SD = 6.69$) and the size of the *high school mathematics package* as an indication of the mathematics level ($M = 5.00$; $SD = 1.75$). As we are aware the self-reports cannot be directly verified, it remained to be seen to which extent GPA and the high school mathematics package were predictive of individual student results. However, given the previously reported positive results in literature for study antecedents towards study success (see introduction), we deemed it reasonable to add both measures to the pool of predictors.

For cognitive ability (with scores ranging from 0 to 20), we tested students on *vocabulary* ($M = 17.59$, $SD = 1.66$, Chronbach's $\alpha = .79$), *comprehensive reading* ($M = 14.90$, $SD = 4.60$, Chronbach's $\alpha = .65$), *mathematics* (for the normal test, $M = 16.53$, $SD = 2.50$, Chronbach's $\alpha = .69$ and for the advanced test, $M = 11.70$, $SD = 3.95$, Chronbach's $\alpha = .96$), *chemistry* ($M = 15.29$, $SD = 2.98$, Chronbach's $\alpha = .98$) and *physics* ($M = 11.88$, $SD = 3.53$,

Chronbach's $\alpha = .96$). For vocabulary, we used the lexTALE test (Lemhöfer & Broersma, 2012) in which students had to assess if the presented stimulus was an existing word or not (60 items). For comprehensive reading, students were asked five questions on a text of medium length about a social psychological experiment. The test had a multiple choice (MC) format with four options. The comprehensive reading test was not administered to students in programs 7, 11 and 15. For the normal mathematics test administered to students in programs 1 to 12 and 20, students had to fill out 20 questions (MC format with four options and open questions) on elementary mathematics. Items included simple math problems like "a book that is on a 40% discount costs \$18. How much did it cost prior to the discount?". For the advanced mathematics test administered to students in programs 13 to 19 and 21, students again had to fill out 20 (MC format with four options and open questions) on more advanced mathematics. Items included problems like "present the general equation of a circle with center (-2, 1) and radius 3." For the chemistry test administered to students in programs 7, 10, 11, 12, 13, 15 and 21, students had to fill out 20 questions (MC format with four options). Items included elementary questions like "what is the total number of valence electrons of a sulfur atom?". For the physics test administered to students in programs 7, 10, 11, 12, 15 and 21, students had to fill out 20 questions (MC format with four options). Items included elementary questions like "What is Newton's first law?". To validate these tests as tests of cognitive ability, we performed a confirmatory factor analysis. All items from the different tests had significant loadings on one common factor (cognitive ability), ranging from .53 to .75 (including the normal mathematics items) or .42 to .86 (including the advanced mathematics items).

For *autonomous* and *controlled motivation*, we assessed students using the Self-Regulation Questionnaire (Vansteenkiste et al., 2009). Students had to respond on a scale from 1 to five how much they agreed with statements like "I'm motivated to study this

program because I'm supposed to do this" (controlled motivation, eight items) or like "I'm motivated to study this program because I want to learn new things" (autonomous motivation, eight items). Students were allocated a score ranging from 0 to 20 for both controlled motivation ($M = 8.31$, $SD = 3.13$, Chronbach's $\alpha = .87$) and autonomous motivation ($M = 15.05$, $SD = 2.37$, Chronbach's $\alpha = .86$). To validate our measurement of motivation, a two-factor solution featured loadings for all items on either controlled (ranging between .69 and .78) or autonomous motivation (ranging between .50 and .83).

For vocational interests, we assessed students using the SIMON-I questionnaire (Fonteyne, Wille, et al., 2017). Students had to respond to 152 items (yes or no), each loading on one of the six *RIASEC* scales. Items included professions like "forrester" (loading on the R-scale) or activities like "writing a scientific paper" (loading on the I-scale). Students were scored from 0 to 100 on the R ($M = 19.10$, $SD = 24.12$, Chronbach's $\alpha = .92$), I ($M = 33.99$, $SD = 21.32$, Chronbach's $\alpha = .88$), A ($M = 28.78$, $SD = 25.13$, Chronbach's $\alpha = .92$), S ($M = 35.73$, $SD = 26.20$, Chronbach's $\alpha = .92$), E ($M = 33.54$, $SD = 28.28$, Chronbach's $\alpha = .93$) and C ($M = 21.33$, $SD = 22.96$, Chronbach's $\alpha = .90$) dimensions. To validate our measurement of vocational interests we performed a randomization test of hypothesized order relations (RTOR, for a full discussion, see Tracey & Rounds, 1997). Results revealed a correspondence index of .92 and a significance of $p = .02$, indicating an excellent circular fit for the current data sample. Next, we also derived two measures of PE fit. *Euclidean distance* ($M = 84.84$, $SD = 48.04$) was calculated analogous to Wille and colleagues (2014), using $P / T = 2 \times R + I - A - 2 \times S - E + C$, $D / I = (1.73 \times E + 1.73 \times C - 1.73 \times I - 1.73 \times A)$ and $\text{Euclidean distance} = \text{SQRT} ((\text{student } P / T - \text{study program } P / T)^2 + (\text{student } D / I - \text{study program } D / I)^2)$. Correlation fit ($M = .71$, $SD = .27$) was calculated analogous to Schelfhout and colleagues (2019a) by making the correlation between the *RIASEC* scores of the student and the *RIASEC* scores of the study program. As the present study will adopt a program-

specific approach, all students within a specific program will study in the same environment. As a consequence, the RIASEC profile of the environment is identical for the students studying the same program. Due to this identical environment, it therefore becomes possible to also investigate the influence of the student's individual RIASEC dimensions on study success. For this reason, we have added both the individual dimensions as well as the fit measures to the pool of possible predictors.

For *conscientiousness* ($M = 150.69$, $SD = 19.78$, Chronbach's $\alpha = .88$) we assessed students using the Personality for Professionals Inventory (De Fruyt & Roland, 2010). Students had to respond to 48 items like "I'm sometimes hard to motivate" on a one to five scale (not characteristic at all – very characteristic). To validate our measurement of conscientiousness, factor analysis revealed loadings of all items on a common factor, ranging from .25 to .63.

For *metacognition knowledge* ($M = 13.61$, $SD = 2.07$, Chronbach's $\alpha = .87$) and *regulation* ($M = 12.99$, $SD = 1.96$, Chronbach's $\alpha = .92$), we assessed students using the Metacognitive Awareness Inventory (Schraw & Dennison, 1994). Students had to indicate on a one to six scale (completely disagree – completely agree) to which degree they agreed to 52 statements (17 for knowledge and 35 for regulation). Scores were rescaled to a score between 0 and 20. Items included statements like "I'm good at remembering information". To validate our measurement of metacognition, factor analyses revealed item loadings from .22 to .73 for knowledge and from .20 to .66 for regulation.

For (cognitive) *test anxiety* ($M = 10.02$, $SD = 2.46$, Chronbach's $\alpha = .92$), we assessed students using the Cognitive Test Anxiety Scale (Cassady & Finch, 2015). Students had to indicate on a one to four scale (totally not characteristic for me – totally characteristic for me) how characteristic they considered statements (25) like "I'm not good at taking exams". Total

scores were rescaled to a score between 0 and 20. To validate our measurement of test anxiety, factor analyses revealed item loadings from .37 to .71.

For *self-control* ($M = 12.79$, $SD = 1.81$, Chronbach's $\alpha = .74$), we assessed students using the Brief Self-control Scale (Tangney et al., 2004). Students had to indicate on a one to five scale (totally not agree – totally agree) how much they agreed to statements (13) like “I have difficulty concentrating”. Total scores were rescaled to a score between 0 and 20. To validate our measurement of self-control, factor analyses revealed item loadings from .43 to .71.

For *academic self-efficacy comprehension* ($M = 14.76$, $SD = 1.62$, Chronbach's $\alpha = .80$) and *academic self-efficacy effort* ($M = 15.23$, $SD = 1.86$, Chronbach's $\alpha = .74$), we assessed students using an adaptation of the College Academic Self-Efficacy Scale (Owen & Froman, 1988). Students had to estimate their capability on a one to five scale (not at all capable – fully capable) of coping with situations or tasks (fourteen comprehension items and eight effort items) like “taking multiple choice exams”. Total scores were rescaled to a score between 0 and 20. To validate our measurement of academic self-efficacy, factor analyses revealed item loadings from .42 to .63 on comprehension and .42 to .71 on effort.

Chapter 5. The Effects of Vocational Interest on Study Results: Student Person–Environment Fit and Program Interest Diversity.

Published online, 4th of April 2019.

*Stijn Schelshout, Bart Wille, Lot Fonteyne, Elisabeth Roels, Filip De Fruyt &
Wouter Duyck (2019).*

*The Effects of Vocational Interest on Study Results:
Student Person–Environment Fit and Program Interest Diversity.
PLoS ONE 14(4), DOI: 10.1371/journal.pone.0214618*

5.1. Abstract

The extent to which a good person-environment (PE) interest fit between student and study program leads to better study results in higher education is an ongoing debate wherein the role of the study program environment has remained inadequately studied. Unanswered questions include: how diverse study programs are in the interests of their student populations, and how this program interest diversity influences study results, in comparison to individual PE fit? The present study addressed these questions in students ($N = 4,635$) enrolled in open-access university education. In such an open access system, students are allowed to make study choices without prior limitations based on previous achievement or high stakes testing. Starting from the homogeneity assumption applied to this open access setting, we propose several hypotheses regarding program interest diversity, motivation, student-program interest fit, and study results. Furthermore, we applied a method of measuring interest diversity based on an existing measure of correlational person-environment fit. Results indicated that interest diversity in an open access study environment was low across study programs. Results also showed the variance present in program interest diversity was linked to autonomous and controlled motivation in the programs' student populations. Finally, program interest diversity better explained study results than individual student fit with their program of choice. Indeed, program interest diversity explained up to 44% of the variance in the average program's study results while individual student-program fit hardly predicted study success at all. Educational policy makers should therefore be aware of the importance of both interest fit and interest diversity during the process of study orientation.

5.2. Introduction

Literature has shown that students who choose a study program that fits their vocational interests, arguably have better study results and have a better chance of finishing higher education in a timely fashion (Tracey & Robbins, 2006). Such study results are usually investigated in settings without much emphasis on the educational access policy. However, these policies show large variety: they can be open or restricted, based on past secondary school performance or tests like the SAT (scholastic aptitude test). This person-environment fit (PE fit) research line focuses heavily on the student side of the PE relationship, while leaving the environment of study programs underexplored (Nauta, 2010). As a consequence, it is still unknown how diverse study environments actually are in terms of student vocational interest, and whether this varies as a function of access policy. For instance, does an academic bachelor in psychology only attract students that are profoundly interested in psychology or do students who enroll in psychology display quite some diversity in their interest pattern? And how does interest diversity in a psychology program compare to the diversity in other programs like mathematics or economics? As literature is still oblivious of study program interest diversity, we are also unsure how this program diversity directly influences study results. In an open access educational environment, the present study uses homogeneity theory and the properties of vocational interest to pose and answer three main research questions. How diverse are student interests within and between study programs? Does this interest diversity directly influence average study results in study programs? And finally, how does this effect of program interest diversity on study results compare to the effects of individual PE fit?

The properties of Vocational Interest

Vocational interest is typically defined as the liking or disliking of certain activities or environments, usually represented by a concise number of dimensions (Lounsbury, Studham,

Steel, Gibson, & Drost, 2009). Vocational interest also has a number of key characteristics which are important when exploring the relation between student interest and study results (Rounds, 1995; Su, Rounds, & Armstrong, 2009).

First, vocational interest has predictive power towards study program choice (Whitney, 1969). Up to 70% of the students (depending on the methodology used) chooses a study program that can be predicted through vocational interest (Burns, 2014; Donnay, 1997). As such, a vocational interest model should be able to compare student interest profile and study program environment on a commensurate scale to explore how good students match their study choice (Tracey, Allen, & Robbins, 2012). For the present study, we have used the RIASEC model by Holland (1997).

Next, student interests always have an object or an environment (Rounds & Su, 2014). As an example, a student can be interested in solving equations or working in an engineering environment. As a consequence, questionnaires targeting higher education students should focus on appropriate items like activities or (future) occupations. For this specific study, we have used the SIMON-I questionnaire (Fonteyne, Wille, Duyck, & De Fruyt, 2017), but our rationale could be easily applied to any other RIASEC-based instrument.

Literature also reports that vocational interests are stable constructs (Swanson & Hansen, 1988). This opens up research possibilities towards prospective studies. As an example, in the present study we have combined students' interests and study results spanning an entire academic year.

Finally, interests are also linked to motivation (Lent, Brown, & Hackett, 1994). Performing actions which the student is highly interested in, like solving equations, can create a study environment that motivates the student towards obtaining his or her degree of choice through facilitating study behavior. Because motivation and vocational interest are linked together, we expect that average program motivation scores of student populations have also

been linked to the interest diversity in student populations. As such, we have also assessed both controlled and autonomous motivation in the present study. Autonomous behavior is performed out of interest or personal importance, while controlled behavior is driven by external demands (Ryan & Deci, 2000). For instance, a student can get good grades because he likes studying the courses in his program of choice (autonomous motivation) or because he wants to adhere to his parents' expectations (controlled motivation).

RIASEC Theory

The Holland RIASEC model has a long standing tradition as one of the most influential models of vocational interest (Holland, 1959, Nauta, 2010; Toomey, Levinson, & Palmer, 2009). The RIASEC theory's basic principle is very straightforward. Persons (students) and environments (study programs) are represented on the same clockwise hexagon, containing six dimensions or scales, using the same RIASEC code: *Realistic*, *Investigative*, *Artistic*, *Social*, *Enterprising*, and *Conventional*. By comparing the profiles of a student and a study program, PE fit indicates how well a (future) student's interests match a specific study program. The current iteration of the RIASEC model still instigates a lot of research and applications specifically targeting (higher) education. As an example, the current study builds on the SIMON-I RIASEC questionnaire that was recently introduced as a study orientation tool, tailored for the transition towards higher education (Fonteyne et al., 2017). SIMON-I measures the interests of (future) higher education students in order to guide the student towards the best fitting study programs. As such, SIMON-I builds on the object characteristic of vocational interest by including items that describe professions (66) and activities (87), tied to one of the six scales. As an example, *Geneticist?* (I) or *Starting up an enterprise?* (E) are two of a total of 153 very short items to which the (future) student had to answer with yes or no. The scores on all dimensions can be recalibrated to a score on a 0-100

scale for each dimension, effectively rendering a student RIASEC profile, e.g. R: 80, I: 70, A:60, S:50, E:55, C:59.

Beside these person profiles, there are a number of ways to construct RIASEC study program environment profiles. One of these methods is built on a common principle that the environment is determined through the people that are in it (Astin & Holland, 1961; Linton, 1945; Schneider, 1987). As vocational interest is a stable construct, students having a good fit with their study programs in year one will likely still have a good fit when they finish their study program. By using successful and persistent students as representatives or *incumbents* of a study program, this incumbent method can empirically generate environment profiles using the profiles of said students (Allen & Robbins, 2010; Holland, 1997). Due to this empirical base, the incumbent method is immune to rater bias, in contrast to profile generation based on expert ratings. Practically, the RIASEC profile of the study program is constructed by averaging out the scores on the RIASEC dimensions using the incumbent student RIASEC profiles of that program (Allen & Robbins, 2010). The relation or (dis)similarity between the individual student and the study program environment is depicted through a measure of PE fit.

Measure of PE Fit: Pearson's Product Moment Correlation Coefficient

Correlational fit is a pattern based similarity measure, that correlates commensurate student and study program RIASEC dimension profile patterns by using the *Pearson's product-moment correlation coefficient*. As such, correlational fit indicates how well the student's interests fit with his or her study program of choice. For instance, fitting a study program profile (M) R:80, I:70, A:60, S:50, E:60, C:70 to a student profile (S) R:80, I:60, A:60, S:50, E:50, C:60 results in a correlation of .87 (or vice versa). A high correlation means a better PE fit. Literature has shown that correlational fit also has predictive value towards study results, especially for first year students (Allen & Robbins, 2010; Tracey et al., 2012).

Although the predictive validity of PE fit between students and their study programs on study results has repeatedly been established, the positive influence of this PE fit on study results remains somewhat limited, varying from a very small to medium effect at best (Assouline & Meir, 1987; Spokane, Meir, & Catalano, 2000; Tinsley, 2000; Tsabari, Tziner, & Meir, 2005). Indeed, some studies report at best a very modest effect of PE fit on study results, while other studies report correlations of up to .32 between study results and PE fit, implying an explained variance of about 10% (Nye, Su, Rounds, & Drasgow, 2012). As a consequence, a debate is still ongoing whether or not the magnitude of these results are in line with the high(er) expectations of theoretical vocational interest models like the Holland RIASEC hexagon (Holland, 1997).

However, when starting a new study related to concepts like PE fit, researchers are usually confronted with a selection bias problem. Access to study programs, especially in the United States, is often already restricted and linked to performance criteria like passing a specific exam or satisfying grade point average (GPA) requirements. As such, these access restrictions yield a pre-selected sample, possibly not only biased in terms of intellectual competence, but potentially also biased towards PE fit between students and programs. It is therefore a possibility that the variation in reported PE fit effects reflects – at least partly – the variation in educational access policies. The present study provides a unique opportunity to enhance our knowledge on the range of PE fit effects by using a large student data set from an open access higher education environment. It will be very interesting to observe whether the positive effect of PE fit on study results generalizes to, or is even stronger under conditions of free study choice, while also examining the influence of program interest diversity on average program results.

Interest Diversity Theory

The age old homogeneity assumption states that people who display similarity in characteristics such as vocational interests tend to lean towards similar environments, with literature spanning more than half a century (Holland, 1966; King et al., 2017; Schneider, 1987). As a consequence, environments like professional occupations are inhabited with individuals that have similar patterns of vocational interests. This internal similarity in the population of an environment also seems to hold in higher education when observing the vocational interest of students. As an example, students that show a good fit between their RIASEC profile and a specific study program are more likely to enroll for such a program than students who do not fit the program (Rounds & Su, 2014; Whitney, 1969). Starting from this internal similarity in the population of study programs, we can make a number of predictions. These predictions form the basis for the present study's research questions.

To start, study programs should display a low interest diversity in their student population. Indeed, individual students who have a high PE fit with a study program are more likely to enroll for this specific study program than students who lack a high PE fit. This mechanism will be even more explicit if there are no further requirements (like GPA or exams) to enter a program as is the case in our present study with open access. As a consequence, the average PE fit in our open access study program population should be quite high when compared to programs with a more restricted access. Due to this high individual PE fit across students, each study program should display high internal similarity, or a low diversity, regarding the vocational interest of its student population. In order to be able and test our predictions of open access versus restricted access, the present study also features data from a small control group (same university) that had to pass an entry exam to enter the Medicine or Dentistry programs.

Next, a high interest in a specific program is just one, autonomous motivation why students enroll for that program³. Students can also opt to act on exterior, more controlled motives to make their choice, like pleasing their parents. As these students are less interested in the specific program, they will have a lower PE fit with the chosen program. Some study programs might be more prone to such externally controlled study choice than others; these programs will attract more students with a lower PE fit and thus show a higher interest diversity in their student population. For the present study, we therefore predict that program interest diversity will vary over programs. We also predict this variance will be linked to autonomous and controlled motivation.

Finally, the internal similarity of the environment will exert an influence on the behavior of the population, with different programs rewarding different interest patterns (Smart, Feldman, & Ethington, 2000). In case of high internal similarity, the behavior of the students will be less determined by their interest pattern but all the more by the study environment. For instance, Tracey and colleagues (2012) used the profile deviation around the mean of the six RIASEC dimensions of a study program as a measure of so called constraint. Smaller deviations represent higher constraint and higher internal similarity. Relevant for this study, results showed a cross-level interaction between program constraint and student interest pattern. Indeed, high program constraint reduced the effect of individual PE fit on study outcomes like GPA and persistence. However, their study did not specifically focus on the direct influence of the environment on study results at the program level, nor on the comparison between the effects of the environment (internal similarity) and the effects of the individual (student PE fit). Our study aims to add to this knowledge. Considering the already hypothesized low levels of program diversity (high internal similarity), we predict a strong effect of the environment for the present study.

³ For a more elaborate discussion on autonomous and controlled motivation, see Vansteenkiste et al., 2009.

Measuring Interest Diversity

The debate regarding the best fitting (dichotomous) measure for studying homogeneity or internal similarity of environments is still undecided. However, Bradley-Geist and Landis (2012) provide evidence that first and foremost, the measures used should depend on the study's major hypotheses. For the present study, the emphasis lies on study program interest diversity and its influence on study results in higher education. As such, interest diversity should be assessable within study programs. Moreover, a program interest diversity assessment should also allow for comparisons across programs so we can investigate the possible influence on study results. For instance, the average deviation measure for testing environment homogeneity reflects the average difference within one group between the individual scores and the mean or median group score (Burke & Dunlap, 2002; Burke Finkelstein, & Dusig, 1999). Regarding vocational interest, we are faced with a conceptual problem of measuring such an average deviation on a single scale. Indeed, the most dominant theories use multiple scales to measure vocational interest. For example, our Holland RIASEC theory uses a 6 dimensional hexagon to depict the vocational interest of individuals and their vocational environments like study programs (see above). As it stands, correlational fit already provides us with a one-scale, parsimonious and continuous PE fit measure that indicates the degree of fit of an individual's interests with his or her environment. Furthermore, a PE fit correlation is already known to be predictive of first year study results (Tracey et al., 2012). Apart from indicating a degree of fit, the correlational fit measure also indicates how far a student's interests deviate from their study program interest profile. By averaging out these deviance measurements over students of a specific program, a continuous measure of program environment interest diversity can be obtained. Such a measure allows us to investigate whether study program populations have a high internal similarity in vocational interest, while still allowing for variance in this hypothesized low program interest diversity.

Present Study

In the present prospective study, we have derived and investigated a number of research questions from the theoretical predictions made in the introduction regarding program interest diversity in an open access environment. For our first question, we have investigated how diverse study programs actually are in the vocational interest of their student population. Students with similar interest patterns should be attracted to similar programs, especially in an open access environment. Consequentially, the fit between students and their program of choice should be high. As such, we hypothesize program interest diversity will be low. We also expect that this general low interest diversity will still show variance over the range of programs, linked to the motivation of their student populations. Indeed, some students will choose the program autonomously, because they are interested in the program itself, while others will take into account exterior motives like parental approval and are less interested in the program itself. We expect some programs could be more prone to such exterior motives of student choice. Therefore, the student population of these programs should show more diverse interest patterns. In sum, we hypothesize that a student population with high autonomous motivation is related to a low program interest diversity, while a population with high controlled motivation is linked to a higher program interest diversity.

For our second question, we have explored if and how program interest diversity has an effect on (average) program study results. On an individual level, literature already showed us that a higher PE fit will lead to better study results (Tracey et al., 2012). One could make the seemingly plausible hypothesis that the program level interest diversity variable derived from this PE fit would show a similar effect, i.e. a lower program interest diversity (high average PE fit of the students) would lead to better results. However, we are wary of making that hypothesis, as we are aware of the ecological fallacy phenomenon that warns researchers to not naively assume that individual effects automatically generalize towards a higher

(program) level (Lubinski & Humphreys, 1996). Moreover, Smart and colleagues (2000) already pointed out that different study programs could reward different interest patterns. As such, we have taken a conservative approach by pitting three hypotheses against each other. We hypothesize rising program interest diversity could show a linear positive effect, a linear negative effect or a curvilinear mixed effect on average program study results. Our findings will also serve as a baseline to integrate possible program interest diversity effects into our third and final research question.

For our final research question, we have compared the found effects of program interest diversity on study results to the effect of individual PE fit. As we are still unsure about the nature of the program interest diversity effect, we can only make predictions regarding individual PE fit. From theory discussed in the introduction, we hypothesize that the individual effects of PE fit on study results will be small, as program environments with low interest diversity (high similarity) are expected to limit the effect of individual PE fit.

5.3. Method and Materials

Data and Procedure

All students ($N_0 = 6,772$, 55% female) starting an academic bachelor at a large Western European university (ranked in the Shanghai top 100, see also www.shanghairanking.com) across eleven faculties and 41 study programs, with an open access policy (anyone who completed secondary education) were invited to participate in a long term assessment to enhance study choice and study results. Though the programs have an open access structure, they are also strictly stratified. In other words, within one program, everyone has to take the same set of study courses during the first year. The study programs are listed in Table 1. Students were asked to participate in the present study during the starting week (end of September 2016) of their curriculum via their lecturers, email and the online learning platform (Fonteyne et al., 2017). Response rate was 71% ($N = 4,827$, 57% female).

Participating students immediately filled out online interest and motivation questionnaires (about five to ten minutes long, see *measures* section for a detailed description). All students were also subject to periodic evaluation systems (once for each course) split up into two sessions (January and May/June 2017) and a retry session (August/September 2017) if they failed the test on the first attempt. At the end of the academic year (September 2017), the results from the SIMON-I test were cross-referenced to the exam results. A total of 4,422 students (92% of N) at least participated in some form of evaluation. The remaining 8% dropped out, prior to any form of evaluation. As literature shows that PE fit also has an influence on perseverance, we took a conservative approach and included the dropouts in our study (Allen & Robbins, 2010). Students from the study programs Medicine and Dentistry ($n = 192$, response rate 86%) already had to pass an entry exam to be allowed to start, in contrast to the students who picked any of the other 39 study programs. As such a high stakes access mechanism (possibly) influences the homogeneity inside a such a program, we have decided to again act conservatively and exclude both programs from our study. However, we have pooled the students from both programs (program RIASEC profiles correlate 0.99 and have about 50% common courses) as a control group for our first question regarding program interest diversity. The final pool of student participants (SP) was $N = 4,635$, spread out across 39 study programs.

Table 1. Study Programs Included in the Present Study.

Number	Programs
1	Psychology
2	Communication Sciences
3	Mathematics
4	Educational Sciences
5	Political Sciences
6	Law
7	Sociology
8	Criminological Sciences
9	Speech Language and Hearing Sciences
10	Physical Education and Movement Sciences

11	Philosophy
12	Linguistics and Literature
13	East European Languages and Cultures
14	History
15	Oriental Languages and Cultures
16	Moral Sciences
17	Art History
18	Archaeology
19	African Studies
20	Veterinary Medicine
21	Physical Therapy and Motor Rehabilitation
22	Pharmaceutical Sciences
23	Bioscience Engineering
24	Economics
25	Biomedical Sciences
26	Engineering – Architecture
27	Engineering
28	Business Economics
29	Bioscience Engineering Technology
30	Engineering Technology
31	Applied Language Studies
32	Biochemistry and Biotechnology
33	Biology
34	Chemistry
35	Physics and Astronomy
36	Geology
37	Geography and Geomatics
38	Computer Sciences
39	Public Administration and Management
	Medicine and Dentistry (excluded)

Besides this first year student pool, we established the profiles of the study programs using interest questionnaire responses of 6,572 senior 3rd and 4th year students spread out over the same 39 study programs featuring in Table 1. These students all met the conditions of academic success and perseverance. The procedure of making the program E-profiles was identical to the procedure used by Allen and Robbins (2010): the RIASEC scores of all students in each study program were averaged out across all six dimensions to obtain the E-profile for each study program.

Measures

Vocational interest

The SIMON-I questionnaire (see also APPENDIX A) was presented to all students to measure the six RIASEC dimensions of the Holland model as described in the introduction (Fonteyne et al., 2017). Derived from the specific item scores, all student participants received a RIASEC profile, with each dimension scoring between 0 (low) and 100 (high). For the present study, the RIASEC scales showed a reliability of .92, .87, .91, .92, .93 and .90 respectively, measured through a Cronbach's alpha (Cronbach, 1951). To confirm the validity of the RIASEC model for the present study's data, we performed a confirmatory factor analysis on the circular structure of the RIASEC dimensions using the CirCe package in R (Browne, 1992; Grassi, Lucio, & Di Blass, 2010). Figure 1 shows the resulting circular structure for the present data sample. Confirmatory factor analysis (CFA) included several measures of model fit (*Standardized Root Mean square Residual* = 0.051; *Normed Fit Index* = 0.97; *Comparative Fit Index* = 0.97; *Goodness of Fit Index* = 0.99), all pointing towards a good to excellent circular fit. For an overview on interpretation of these indices, we refer to the exhaustive listings provided by Kenny (2015). Additionally, we also verified the circular RIASEC order and structure by conducting a *randomization test of hypothesized order relations* (RTOR) using the RANDALL package (Hubert & Arabie, 1987; Tracey, 1997; Tracey & Rounds, 1996). Results of this RTOR analysis revealed a *correspondence index* of .92, while a circular fit of the data also reached significance, $p = .02$. For a full discussion on the RANDALL function and RTOR analyses, we refer to (Hubert & Arabie, 1987; Tracey, 1997; Tracey & Rounds, 1996). In sum, the theorized circular fit for our RIASEC data was confirmed by both CFA and RTOR analyses.

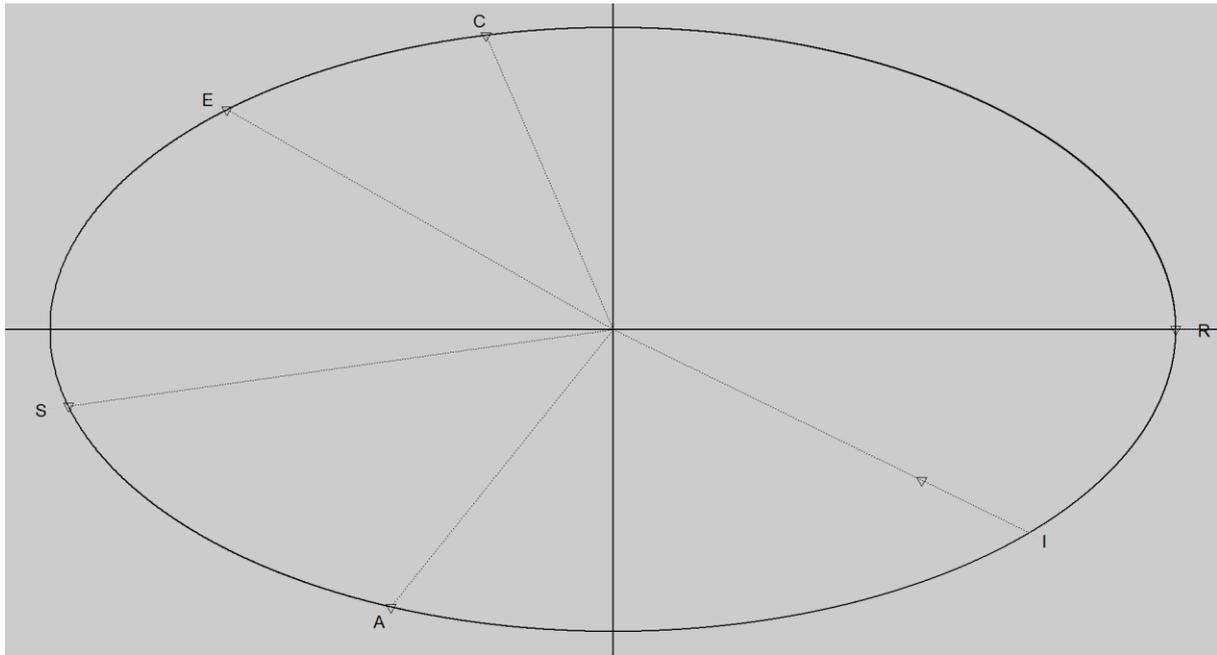


Figure 1. Verified Study Specific Circumplex Structure of the RIASEC Model. R = realistic dimension, I = investigative dimension, A = artistic dimension, S = social dimension, E = enterprising dimension, C = conventional dimension.

Student PE interest fit and study program interest diversity

Next, the student PE fit between the vocational interest of the student and her/his study program was established using the correlational fit measure. To calculate the correlational fit measure (shape resemblance between the hexagonal pattern of student and study program profile) for each student, each student's RIASEC profile was correlated with his or her study program RIASEC profile based on the profiles of successful and persistent students. As this correlational fit measure also represents a deviance, the measure (on a scale of -1 to 1) was then rescaled to an easy-to-interpret interest deviance between 0 and 1, with $D\text{ COR} = 1 - (\text{correlational fit} + 1) / 2$. As an example, a student with a correlational fit measurement of 0.76 would rescale to $D\text{ COR} = 0.12$, indicating the student's RIASEC profile deviates 12% in regard to the profile of his or her study program.

A program interest diversity measure should indicate how deviant students' interests are in a given study program of choice. By averaging out D COR across students in a program, we can obtain a measure of interest diversity AD COR for each study program, that

represent how much students diverge on average from the program profile. As an example, a study program with an AD COR of 0.23 indicates that the RIASEC profiles of students within the program deviate (on average) 23% in regard to the program profile.

Autonomous and controlled motivation

The Self-Regulation Questionnaire was presented to all students to measure their autonomous (8 items) and controlled (8 items) motivation (Vansteenkiste, Sierens, Soenens, Luyckx, & Lens, 2009). For the present study, the autonomous and controlled motivation subscales showed a reliability of .86 and .87 respectively, measured through a Cronbach's alpha (Cronbach, 1951). A factor analyses on both the autonomous and controlled motivation subscales revealed the expected two factor structure, explaining 52% of the variance. Items from the autonomous subscale showed high loadings on the autonomous factor ($M = 0.70$, ranging from 0.50 to 0.83) but not on the controlled factor. In contrast, items from the controlled subscale showed high loadings on the controlled factor ($M = 0.73$, ranging from 0.70 to 0.77).

Study results

GPA is a widely known and used measure of study success (Richardson, Abraham, & Bond, 2012). However, as Graham (2015) already pointed out, the validity and reliability of such a measure cannot be merely assumed, which forms a widespread problem in literature concerning study success. Indeed, validity and reliability are function of both sample and measure, rendering GPA results from past samples insufficient (Harris, 2003). To ensure reliability and validity of the GPA measure, not only towards research, but especially towards the eventual degree of students, the present study's featuring open access university has installed several precautions embedded in the teaching and grading procedures for each study program. Considering the open access, it is important to note that all programs are strictly stratified in the first year. Because all students take the same courses, GPA is fully

comparable within a program. For means of validity, attainment levels for all programs are actively protected by national and regional law. In other words, what students need to know in theory and practice is officially decreed and controlled by the government. For means of reliability, each program consists of 60 study points, divided over several courses (about ten courses for each program), taught by different professors and lecturers to avoid rater bias. The exams for each program's course are spread across a number of (non) periodical methods including but not limited to written exams, multiple choice exams (usually corrected automatically by use of computer), oral exams and essay writing to avoid methodical bias. In case of failing an exam, a student always has the right to resit the exam and even an appeal if irregularities (like dubious questions) were established. Both the resit and the appeal actively counter low reliability of exams. Importantly, we have included a PASS measure in the present study. In order to earn a PASS for his or her first year (and all other years), a student has to pass all courses of a study program by obtaining a course score of at least 10/20 for each course. This passing measure is an excellent countermeasure against possible overly optimistic exam marking in some courses of a program, as a student has to pass all courses of that program to obtain a PASS. Indeed, if a student earns a PASS, the results from all courses unanimously indicate the student has mastered the learning material for the first year. On the other hand, if there are courses in a program that are too difficult in comparison to the other courses, to the extent that no one would succeed the program, the discrepancy between the PASS rate (which would be zero) and the average GPA of the program should alert us to problematic (too strict) grading for those specific courses. For these specific reasons, we have included both PASS and GPA as measures of study results. GPA indicates the global result of the student in his first year study program on a scale from 0 to 1,000. PASS was assessed dichotomously (1/0) at the first year student level, and averaged out at the program level as a PASS rate, across students. If these study result measures are reliable and valid, both

measures should correlate highly and should display similar results at the program level.

Despite these precautions, we still deem it possible (although unlikely) that a student's GPA could be biased due to specific program choice and subsequent deviant grading. As such, we have taken a conservative approach by obtaining an estimate of the maximal hypothetical bias for GPA by calculating the intra class correlation coefficient through a multilevel model. For our purposes, this coefficient indicates how much of the variance in individual GPA can be attributed to the program level. Practically, the coefficient provides us with an estimate of the maximum bias due to possible different grading standards in different programs. However, stronger students could also systematically choose more difficult programs. This confound cannot be disentangled within the scope of the present study. As a result, we have included the intra class correlation coefficient as an absolute maximum of hypothetical bias in a student's GPA due to the program of his/her choice.

Analyses

To avoid unwanted bias of (potentially) skewed distributions of the correlational data on outcomes and subsequent conclusions, we performed all statistical testing on standardized scores by taking the z - scores of the D COR measures (by subtracting the grand mean and dividing by the grand standard deviation of the full data set) as the base measure. Averaging out these z - scores over programs renders the standardized equivalent of the AD COR measures.

To address our first research question, we inspected the variance of interest diversity over programs. The interest diversity of open access programs was tested against our control group that had to pass an entry exam. We also regress interest diversity (AD COR) on autonomous and controlled motivation to test whether both types of motivation in the student population are indeed linked to the variance in program interest diversity. To address our second research question regarding the influence of interest diversity on study results, we

have regressed the average study program GPA and PASS on program interest diversity. As we are pitting three hypotheses against each other, we have considered both linear and curvilinear relations. To address our final research question regarding the comparison between the individual PE fit effect and the environmental program interest diversity effect, we have constructed two hierarchical models, a linear one for GPA and a logistic one for PASS. In these models, we have investigated the effect of individual PE fit on study results, the effect of program interest diversity on study results and the cross-level interaction between individual PE fit and study program diversity. All analyses were conducted using R(Studio), SPSS and HLM software.

5.4. Results

Table 2 shows the program summary containing the reference number, the scores on the R, I, A, S, E and C dimensions, the average student GPA, the student PASS rates, the AD COR interest diversity, the scores on autonomous (AMOT) and controlled (CMOT) motivation, the response rate (RR) on the SIMON-I questionnaire and finally the number of students (N). The correlation of GPA and PASS rates amounted to 0.80. Closer inspection revealed there was one program, *Geology* (36) with 0% PASS rate. Because this find could prove problematic, we considered the average GPA (370.29) of the program, which was very low (compared to the other programs). The correlation between GPA and PASS did not change by excluding *Geology* (36). There does not seem to be a huge discrepancy between the GPA and PASS results for *Geology* (36). Because 0% PASS rate is still a huge outlier, we decided to act conservatively and pay special attention to the *Geology* (36) result in specific PASS analyses.

Table 2. Descriptive Statistics for all Study Programs.

Number	R	I	A	S	E	C	GPA	PASS	AD COR	AMOT	CMOT	N	RR
1	6.36	32.06	35.97	58.17	27.10	10.45	521.28	0.49	0.12	15.61	7.94	424	0.82
2	8.52	22.68	56.67	36.72	55.67	14.97	414.40	0.31	0.11	14.7	8.06	121	0.95

3	19.65	37.31	15.89	14.31	18.51	27.47	512.23	0.50	0.26	15.27	8.35	26	0.67
4	4.08	24.21	33.78	74.95	25.05	13.76	560.56	0.67	0.06	15.16	8.20	106	0.88
5	9.50	25.53	38.85	37.69	53.31	20.15	418.59	0.34	0.18	15.02	8.48	86	0.90
6	7.56	25.10	32.02	37.93	55.87	35.51	407.08	0.27	0.16	15.39	8.13	298	0.61
7	6.15	30.46	42.92	50.85	37.97	11.83	390.17	0.23	0.13	15.99	8.37	35	0.85
8	9.87	27.75	27.88	46.99	30.12	19.19	461.54	0.33	0.21	15.1	8.21	149	0.60
9	6.13	34.34	37.18	65.38	23.28	12.43	633.17	0.60	0.09	15.09	8.70	60	0.91
10	8.18	29.01	20.26	37.59	25.12	13.46	436.94	0.35	0.22	14.84	8.00	49	0.70
11	7.13	34.41	56.02	45.76	31.17	17.36	408.00	0.50	0.13	14.75	7.00	6	0.29
12	6.86	22.71	59.15	38.71	21.47	7.43	484.55	0.45	0.10	15.39	8.77	161	0.83
13	4.97	24.24	57.69	51.44	29.44	7.51	626.00	0.71	0.06	16.64	6.86	7	0.41
14	11.66	28.55	49.22	31.18	28.81	12.94	453.04	0.33	0.20	15.59	8.26	75	0.64
15	6.25	20.00	43.34	33.79	19.57	8.48	356.73	0.18	0.14	14.42	8.89	33	0.45
16	8.73	41.63	60.18	52.38	27.20	6.55	372.50	0.50	0.12	16.5	7.83	6	0.30
17	10.32	24.33	68.82	34.06	20.10	5.70	339.78	0.19	0.08	16.55	7.90	36	0.49
18	28.71	43.49	43.64	17.83	15.25	10.27	435.29	0.29	0.16	14.26	8.24	17	0.53
19	7.32	31.56	56.32	61.76	25.00	7.97	577.17	0.83	0.05	17.06	5.42	6	0.46
20	18.08	45.68	22.78	31.77	18.06	13.69	428.40	0.34	0.18	15.49	7.36	179	0.66
21	13.79	40.32	21.61	46.63	17.87	11.62	430.13	0.27	0.15	14.81	7.77	361	0.75
22	15.17	49.44	23.51	36.26	20.64	17.74	524.65	0.49	0.17	15.38	8.35	210	0.78
23	32.60	59.13	20.36	24.49	27.10	20.09	479.13	0.41	0.18	15.35	8.93	202	0.81
24	22.03	25.20	21.19	22.84	66.53	54.16	524.12	0.45	0.12	14.74	8.96	360	0.69
25	16.03	54.35	26.30	34.54	20.76	14.99	503.03	0.39	0.14	15.25	7.98	129	0.78
26	46.44	30.82	56.25	20.56	31.20	15.70	486.63	0.48	0.18	15.06	7.86	62	0.43
27	51.95	40.67	22.22	13.61	32.86	22.57	539.51	0.52	0.19	15.19	8.70	222	0.65
28	14.10	16.84	23.73	27.22	66.83	47.84	476.72	0.39	0.12	14.12	8.96	330	0.61
29	33.53	45.37	19.20	20.84	22.78	19.14	524.52	0.45	0.23	13.93	9.28	75	0.82
30	59.70	32.15	22.71	13.39	26.77	19.42	421.68	0.29	0.18	14.09	8.20	349	0.84
31	6.22	16.44	47.33	35.99	30.01	10.94	433.15	0.33	0.14	14.83	8.75	133	0.67
32	17.12	54.26	22.00	22.97	13.05	11.32	463.81	0.37	0.13	14.9	8.12	67	0.66
33	23.24	50.75	28.77	27.18	13.18	6.94	390.70	0.36	0.12	15.06	7.12	50	0.68
34	22.94	47.74	17.11	21.77	18.10	17.70	479.05	0.39	0.18	14.61	7.53	38	0.66
35	33.58	46.81	23.97	11.56	15.11	11.03	400.60	0.40	0.15	15.47	7.71	47	0.68
36	33.90	46.25	31.00	20.11	15.59	10.42	370.29	0.00	0.17	14.54	8.48	14	0.67
37	38.45	39.52	20.57	29.66	13.64	15.69	522.15	0.46	0.26	14.88	10.79	13	0.59
38	31.50	26.76	24.90	7.58	20.09	15.64	459.51	0.51	0.24	14.44	7.73	43	0.66
39	8.74	18.79	25.18	42.15	71.35	42.78	476.58	0.44	0.11	15.01	8.39	50	0.70
*	17.51	47.68	23.41	44.68	26.40	14.81	702.37	0.93	0.17	16.38	8.36	192	0.86

Note. Number: see Table 1, R = average program scores (on 100) on realistic dimension, I = average program scores (on 100) on investigative dimension, A = average program scores (on 100) on artistic dimension, S = average program scores (on 100) on social dimension, E = average program scores (on 100) on enterprising dimension, C = average program scores (on 100) on conventional dimension, GPA = average program grade point average, PASS = average program pass-rate, AD COR = program interest diversity, AMOT = average program autonomous motivation, CMOT = average controlled program motivation, N = number of students in the program, RR = program response rate.

Important to note, the intra class correlation coefficient for GPA amounted to 6%. This

means that only 6% of the variance in GPA can be attributed to the program level. In other words, the individual GPA is determined for 94% based on personal achievement. As such, the maximum possible bias of GPA due to non-equivalent grading can be estimated at 6%.

Research Question 1: Interest Diversity of Study Programs

Figure 2 shows the spread of interest diversity (AD COR) across programs ($M = .15$; $SD = .05$). This interest diversity also corresponds to a rather impressive average PE fit correlation of .70 between the RIASEC profiles of students and their program of choice. We thus clearly observe a concentration of programs at the lower end of the diversity continuum (due to the high average PE fit), ranging from .05 (African Studies) to .26 (Mathematics). This concentration at the lower end also means that 74% of the interest diversity continuum at the higher end remains unused. To formally test if an open access environment indeed results in more interest diversity regarding study programs, we compared the interest diversity of 39 study programs with open access to the control group with restricted access. A two-sided one sample t-test indeed revealed a significant difference, $t(38) = -4.27$, $p < .001$, with a somewhat large effect size of Cohen's $d = 0.69$. We therefore conclude that the evidence shown confirms our hypothesis. Study programs in an open access environment indeed display a low interest diversity regarding their student population.

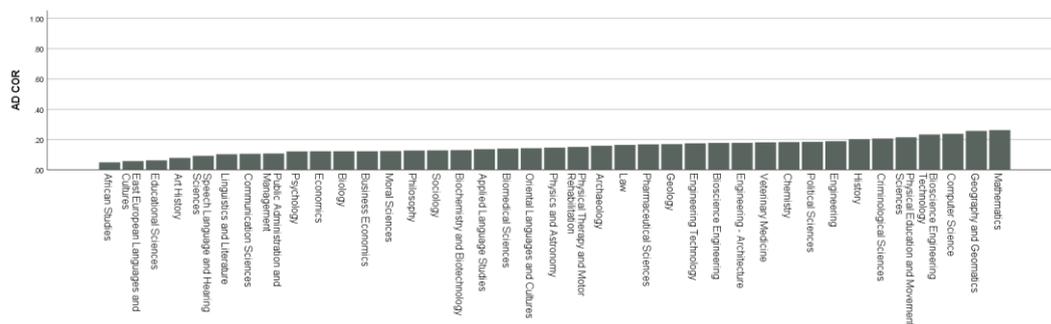


Figure 2. Interest Diversity Program Spread across the 0-1 Continuum. The X-axis displays the 39 study programs featuring in this study. Y-axis displays the program interest diversity expressed through an AD COR value.

For the second part of our first research question, Figure 2 also shows interest diversity does still display quite some variance across programs at the lower end of the

continuum. For instance, the interest diversity in *Mathematics* is about five times larger than the diversity in *African Studies*. To test whether this variance is linked to motivation, we regressed interest diversity (AD COR) on (autonomous and controlled) motivation. The omnibus test proved to be significant, $F(2, 36) = 6.86, p = .003$, with an explained variance of 28%. Standardized regression coefficients ($\beta_1 = -0.33, \beta_2 = 0.28$, for autonomous and controlled motivation respectively) indicate a negative relation between program interest diversity and average autonomous motivation in the program student population, and a positive relation between interest diversity and average controlled motivation in the program student population. Closer graphical inspection (Figures 3 and 4) of the individual effects also reveal the effects are especially present at (relatively) very high and very low levels of controlled and autonomous motivation. These findings confirm our hypothesis that program interest diversity is indeed related to motivation in an open access environment. Programs with low interest diversity have a student population with high autonomous motivation and vice versa. Programs with high diversity are linked to student populations with a higher controlled motivation while programs with low diversity are linked to populations with a lower controlled motivation.

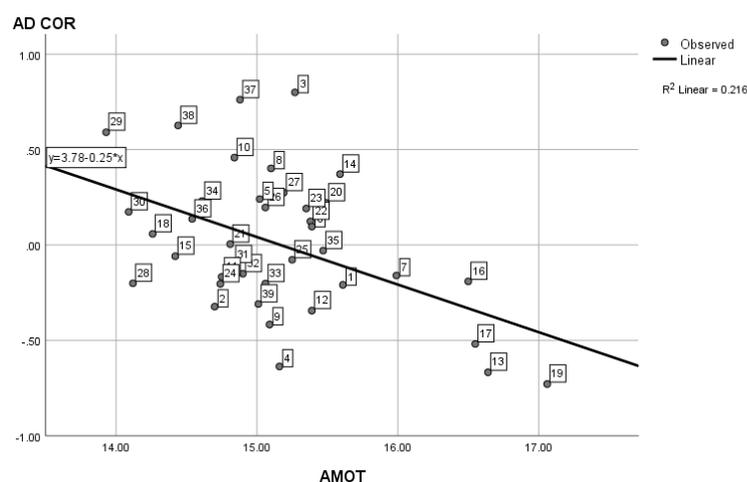


Figure 3. The Linear Regression of Interest Diversity on Autonomous Motivation. Interest Diversity is measured using AD COR (based on D COR z – scores), with AMOT = autonomous motivation. The data points (dotted study programs) are labeled analogous to Table 1.

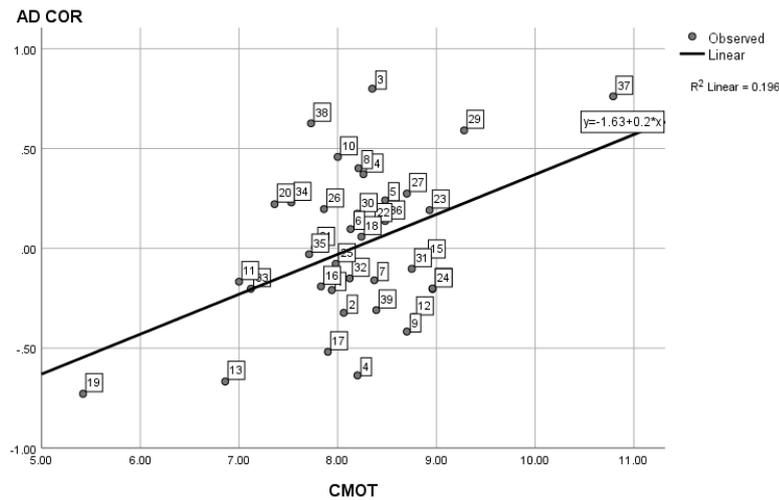


Figure 4. The Linear Regression of Interest Diversity on Controlled Motivation. Interest Diversity is measured using AD COR (based on D COR z – scores), with CMOT = controlled motivation. The data points (dotted study programs) are labeled analogous to Table 1.

Research Question 2: Effects of Interest Program Diversity

To examine the possible direct effect of interest diversity in study programs on average program study results, we conducted two (curvi-) linear regressions of both average program GPA and PASS rates on interest diversity (AD COR). Figures 5 and 6 show the regressions of average program study results on program interest diversity. We obtained similar results for both measures of study results. Indeed, the linear regressions of study results on interest diversity were not significant $F(1,37) = 0.63, p = .43$ and $F(1,36) = 3.01, p = .09$ for GPA and PASS respectively. The curvilinear regressions of study results on interest diversity however, were significant with high levels of explained variance, $F(2,36) = 7.19, p = .002, R^2 = .29$ and $F(2,35) = 13.84, p < .001, R^2 = .44$ for GPA and PASS respectively. Adding the *Geology* (36) results to the PASS regression rendered similar results. These curvilinear findings provide evidence that rising program interest diversity has a mixed effect on the average study results of the program’s student population. Different programs do seem to reward different interest patterns (Smart, 2000). More specifically, a very low diversity in a small number of programs is associated with very high average study results. However, for

the majority of the programs that had a higher diversity to begin with, we observe that study results tend to improve as program interest diversity rises.

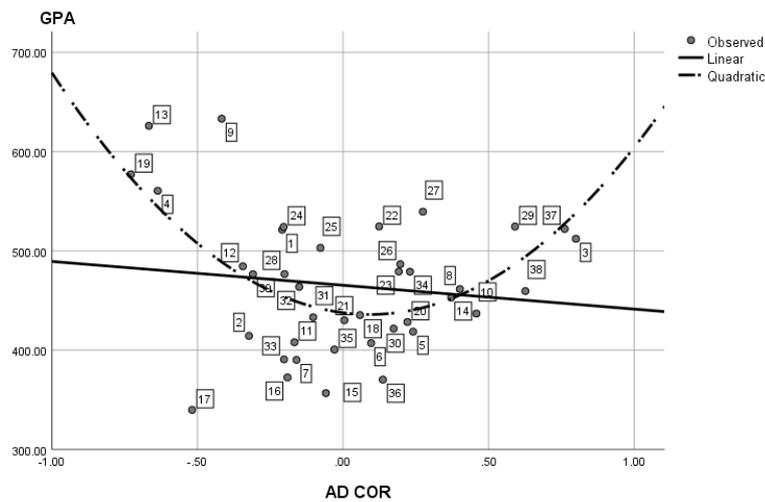


Figure 5. The (Curvi-) Linear Regression of Study Program Average Results (GPA) on Study Program Interest Diversity. Interest Diversity is measured using AD COR (based on D COR z – scores), with GPA = average grade point average for each study program. The linear regression is depicted as a full line, the quadratic regression is depicted as an interrupted line. The data points (dotted study programs) are labeled analogous to Table 1.

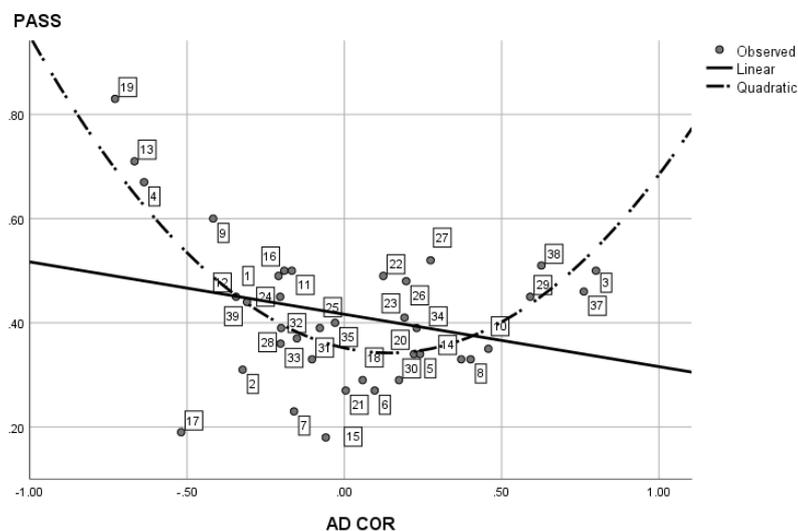


Figure 6. The (Curvi-) Linear Regression of Average Study Program Results (PASS) on Study Program Interest Diversity. Interest Diversity is measured using AD COR (based on D COR z – scores), with PASS = average pass rate for each study program. The linear regression is depicted as a full line, the quadratic regression is depicted as an interrupted line. The data points (dotted study programs) are labeled analogous to Table 1. With a pass rate of 0%, Geology (36) was considered an outlier and was removed from analyses.

As we were curious about the nature of these curvilinear effects, we ran further post hoc analyses. Closer inspection of Figures 5 and 6 revealed that the left, descending part of the curvilinear relation is largely caused by the study results and interest diversity of four

programs: *Educational Sciences, Speech Language and Hearing Sciences, East European Languages and Cultures* and *African Studies*. We decided to compare the six dimension RIASEC profiles of these four programs to each other and to the other programs. Table 3 shows the correlation of the four RIASEC program patterns. The high correlations indicate these programs have very similar interest patterns. When comparing the RIASEC interest profile of these four programs to the other programs in the present study, these specific programs have relatively very high scores on the social S dimension (rankings 1, 2, 3 and 6 out of 39), and very low scores on the realistic R dimension (rankings 30, 37, 38 and 39 out of 39). For these specific programs, a low interest diversity is tied to better study results, as is shown by the correlation between their interest diversity and their average study results: AD COR correlates -.47 with GPA and -.88 with PASS.

Table 3. Comparison (correlations) RIASEC Profiles of *Educational Sciences, Speech Language and Hearing Sciences, East European Languages and Cultures* and *African Studies*.

Number	Program	4	9	13	19
4	Educational Sciences	---	.97	.79	.87
9	Speech Language and Hearing Sciences		---	.83	.93
13	East European Languages and Cultures			---	.97
19	African Studies				---

When considering the other programs of the present study, we repeated the regression of study results (GPA and PASS) on interest diversity by excluding the results from these four programs. Figure 7 now clearly shows a linear regression best explains the relation between program interest diversity and average study results compared to a curvilinear one.

Statistically, the linear regression of GPA on interest diversity reached significance, $F(1, 33) = 6.13, p = .02, R^2 = .16$, while the curvilinear one no longer did, $F(2, 32) = 3.10, p = .06$.

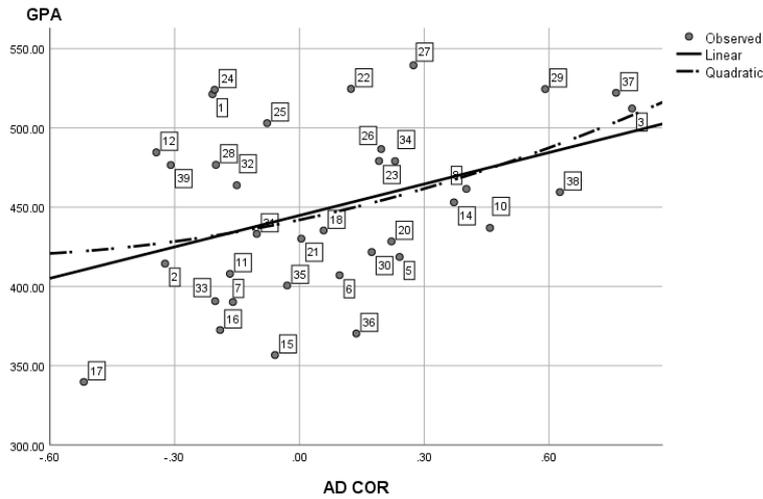


Figure 7. Post Hoc Analysis: The (Curvi-) Linear Regression of Average Study Program Results (GPA) on Study Program Interest Diversity. The data points (study programs) are labeled analogous to Table 1 and are identical to Figure 4, with omission of points 4, 9, 13 and 19 that were discussed separately.

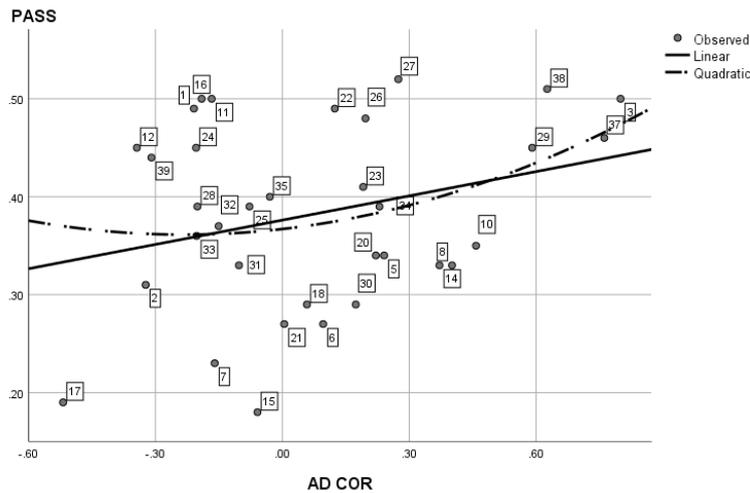


Figure 8. Post Hoc Analysis: The (Curvi-) linear Regression of Average Study Program Results (PASS) on Study Program Interest Diversity. The data points (study programs) are labeled analogous to Table 1 and are identical to Figure 5, with omission of points 4, 9, 13 and 19 that were discussed separately. With a pass rate of 0%, Geology (36) was considered an outlier and was removed from analyses.

Figure 8 shows a somewhat similar pattern. Although the linear regression of PASS on interest diversity did not reach significance, $F(1, 32) = 2.84, p = .10$, interest diversity still showed a strong linear trend towards an effect by explaining 8% of the variance in program study results. The curvilinear regression did not reach significance, $F(2, 31) = 1.72, p = .20$.

The addition of *Geology* (36) to the regression rendered similar results, although the explained linear variance only amounted to 5%.

In sum, these results confirm our hypothesis that interest diversity has a direct, but mixed effect on average study program results in an open access environment. Different programs seem to reward different interest patterns. In general, a higher interest diversity of the student population has a positive effect on average program study results. However, this effect seems to reverse for very specific programs displaying a very high social dimension and a very low realistic dimension. These programs reach high average study results if the student population shows very low levels of interest diversity.

Research Question 3: Individual Student PE fit and Program Interest Diversity

To compare the effect of program interest diversity on study results to the effect of individual PE fit, we have performed multilevel analyses of student (level 1) and program (level 2) effects on study results. As GPA is a linear variable and PASS is a dichotomous variable, we constructed a linear and a logistic multilevel model. Effect sizes for the different effects at the individual level, program level or both levels combined (full model) were calculated by combining the variance components (GPA model) and model deviance statistics (PASS model) into a pseudo R^2 .

Table 4 shows the final version of the GPA model (see also Brambor, Clark, & Golder, 2006 for a discussion on multiplicative interaction analysis). Though significant, the fit of the individual student with this study environment (measured through the D COR measure) only explained 0.6% of the variance in study results at the specific individual level and the general full GPA model. In contrast, program interest diversity (AD COR) explained 17% of the variance in average study results at the program level and 1% of the variance in the full model.

Table 4. Multilevel GPA Model.

A
Multilevel GPA Model
Level-1 Model
$GPA_{ij} = \beta_{0j} + \beta_{1j} * D COR_{ij} + r_{ij}$
Level-2 Model
$\beta_{0j} = \gamma_{00} + \gamma_{01} * AD COR_j + \gamma_{02} * AD COR_j^2 + u_{0j}$
$\beta_{1j} = \gamma_{10} + u_{1j}$
Mixed Model
$GPA_{ij} = \gamma_{00} + \gamma_{01} * AD COR_j + \gamma_{02} * AD COR_j^2$ $+ \gamma_{10} * D COR_{ij}$ $+ u_{0j} + r_{ij}$

B	Coefficient	Standard Error	t-ratio	appr. d.f.	p-value
Final estimation of fixed effects					
Fixed Effect					
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	446.22	10.12	44.07	36	< .001
AD COR, γ_{01}	-28.73	31.63	-0.91	36	0.37
AD COR ² , γ_{02}	178.33	55.06	3.24	36	0.003
For D COR slope, β_1					
INTRCPT2, γ_{10}	-18.03	3.54	-5.09	4595	< .001

C	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Final estimation of variance components					
Random Effect					
INTRCPT1, u_0	48.66	2367.93	36	259.14	< .001
level-1, r	213.16	45440.33	---	---	---

Note. GPA = grade point average, D COR = student PE interest fit, AD COR = program interest diversity, appr. d. f. = approximated degrees of freedom. Model construction was conducted as follows. We first determined the amount of variance in study results (GPA) generated through both levels by using the intercept-only model. Level 1 (individual level) generated 94%, while level 2 (program level) generated the remaining 6% of variance in study results measured through GPA. Explained variance of the different effects was calculated using comparisons to the intercept-only model. Before making the actual models, we tested the possible influence of study program SIMON-I response rate and population (number of students) by adding them to the intercept model. Both tests were not significant, $t(36) = 1.14, p = .26$ and $t(36) = 0.41, p = .69$ respectively. Both group variables were thus removed from the intercept model. The D COR predictor was added to the model as a program centered variable, removing the variance between programs. This variance between programs was then added through the curvilinear effect of AD COR from research question 2 completing the model. Note that the linear term AD COR renders a non-significant result, while the quadratic term did reach significance. Though the absence or presence of the linear term does not change the curvilinear nature of the interest diversity effect, we have decided to keep the linear term as a part of the model due to the multiplicative interaction between the student and the program level.

As hypothesized regarding our third question, our GPA multilevel model thus indicates the influence of the individual's PE fit on individual study results can indeed be considered low. However, at the same time, these GPA model results also reveal that interest diversity over programs explains much more variance in average study results in that program compared to the individual level.

We also tested the addition of a cross-level interaction (through a random slope for PE fit) between interest diversity (AD COR) at the program level and PE-fit (D COR) at the individual level. The chi-squared test returned a non-significant result, $\chi^2(38) = 45.26, p = 0.20$. This result indicates the individual PE fit does not interact with the program interest diversity. In other words, there is no different individual effect of PE fit on study results depending on the study program environment.

Table 5 shows the final PASS model. The results are largely analogous to those from the GPA model. The deviance statistic from the full models (containing predictors) compared to the intercept only model revealed that the predictor models were significant. However, the models showed very low pseudo R^2 (around 0.1%), for both individual PE fit and program interest diversity when considering full model deviance (all levels combined). PE fit only reached a pseudo R^2 of about 0.1% at the individual level, while the pseudo R^2 for interest diversity did reach 37% at the program level.

Table 5. Multilevel PASS Model.

A
Multilevel PASS model
Level-1 Model
$Prob(PASS_{ij} = 1 \beta_j) = \phi_{ij}$
$log[\phi_{ij}/(1 - \phi_{ij})] = \eta_{ij}$
$\eta_{ij} = \beta_{0j} + \beta_{1j}*(D COR_{ij})$
Level-2 Model
$\beta_{0j} = \gamma_{00} + \gamma_{01}*AD COR_j + \gamma_{02}*AD COR_j^2 + u_{0j}$
$\beta_{1j} = \gamma_{10} + u_{1j}$

Mixed Model

$$\eta_{ij} = \gamma_{00} + \gamma_{01} * AD\ COR_j + \gamma_{02} * AD\ COR_j^2 + \gamma_{10} * D\ COR_{ij} + u_{0j}$$

B					
Final estimation of fixed effects	Coefficient	Standard Error	t-ratio	appr. d.f.	p-value
Fixed Effect					
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	-0.38	0.07	-5.44	36	< .001
AD COR, γ_{01}	-0.4	0.23	-1.715	36	0.1
AD COR ² , γ_{02}	1.91	0.48	4.02	36	< .001
For D COR slope, β_1					
INTRCPT2, γ_{10}	-0.15	0.04	-4.1	4595	< .001
C					
Final estimation of variance components	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Random Effect					
INTRCPT1, u_0	0.33	0.11	36	157.01	< .001
Model Deviance = 14571 – 5 estimated parameters					

Note. GPA = grade point average, D COR = student PE interest fit, AD COR = program interest diversity, appr. d. f. = approximated degrees of freedom. Model construction: analogous to the multilevel GPA model. Program SIMON-I response rate and population again tested non-significant, $t(36) = 0.60, p = .55$ and $t(36) = -0.46, p = .65$ and were removed from the model.

As hypothesized regarding our third question, our PASS multilevel model thus indicates the influence of the individual's PE fit on individual study results can indeed be considered low. In contrast to the GPA model, both individual PE fit and program interest diversity displayed very low levels of explained variance when considering full model deviance. Analogous to the GPA model, PASS model results also reveal that interest diversity in the program explains much more variance in typical passing rates for that program, across students.

The addition of a random slope for the PE predictor again resulted in a non-significant result, $\chi^2(38) = 35.04, p > .50$. Hence, we do not find any evidence for a cross-level

interaction between individual student PE fit and program interest diversity on study results.

In sum, these findings again confirm our hypothesis that the individual effect of PE fit on study results is small to almost non-existent in an open access environment, while interest diversity has a profound effect on results at the program level. For the full models, we have found mixed evidence that program interest diversity is more explanative towards study results than individual PE fit. Furthermore, we did not find any evidence for an interaction effect between group level program interest diversity on individual study results.

5.5. Discussion

Vocational interest refers to the liking or disliking of certain activities or environments, represented by a number of base dimensions and characterized by the properties of prediction, contextualization, stability and motivation (Burns, 2014; Ryan & Deci, 2000; Donnay, 1997; Lounsbury et al., 2009; Rounds, 1995; Rounds & Su, 2014; Su et al., 2009; Swanson & Hansen, 1988; Whitney, 1969). Literature has shown that individual person-environment fit (PE fit) between students and study programs influences higher education study results (Allen & Robbins, 2010; Assouline & Meir, 1987; Nye et al., 2012; Spokane et al., 2000; Tinsley, 2000; Tracey et al., 2012; Tracey & Robbins, 2006; Tsabari, Tziner, & Meir, 2005). These studies however, mostly focus on the student, while leaving the study environment underdeveloped (Nauta, 2010). As such, we are uncertain how diverse study programs actually are in terms of vocational interest of their student populations. Moreover, we also do not know whether and how program interest diversity exerts an influence on study outcomes like grade point average (GPA). Finally, when studying theoretical concepts like interest diversity or PE fit, admission restrictions to study programs in higher education, like entry exams or GPA requirements, may imply a selection bias in student intake that may influence the effects of PE fit.

The present prospective study set out to remedy these voids in literature. As such, the present study was conducted in an open access environment, exploring the interactions between individual interest PE fit and environment program diversity. Although we derived our hypotheses from homogeneity theory, we do not consider interest diversity as a mere synonym for homogeneity (Holland, 1966; King et al., 2017; Schneider, 1987). In fact, our operationalization of the interest diversity construct is quite unique, as it reuses measures of PE fit as an indication of how a student deviates from his study program profile. By averaging out these deviances across a program, a continuous measure of program interest diversity was obtained, through the use of a very large sample of students. Using this interest diversity measure, we assessed three research questions. We investigated these questions in a population of bachelor students starting their academic trajectory at a large Western-European university (Shanghai top 100) across eleven faculties with an open access policy.

During the present study, special care was given to the validity and reliability of predictors and study result measurement. As indicated by Graham (2015) and Harris (2003), reliability and validity are function of both sample and measure. The university where the study took place already had a number of measures in place to guard the reliability and validity of study results. Apart from the widely known GPA measure (Richardson et al., 2012), we also added an extra measure of study success through the PASS rate. A student only received a PASS if he or she succeeded for all courses of the program. As explained in the method section, if our study result measures are reliable, both measures should show a high correlation and both measures should show the same result pattern. Our analyses indeed confirmed both predictions. Moreover, the intra class correlation coefficient for GPA amounted to only 6%. This find indicates that the bias in GPA due to non-equivalent quoting can only amount to a maximum of 6%. To which extent this percentage is determined by stronger students systematically choosing certain programs or rater-bias cannot be

disentangled within the current study. Still, as a result, a student's individual GPA is to a very large degree (at least 94%) determined by his personal achievement and not through the specifics of program followed. As such, we are convinced that we have taken the necessary precautions to ensure the reliability and validity of our study result measures and the overall results of the present study.

For our first research question, we investigated how diverse study programs actually are in the vocational interest of their student population. We hypothesized program interest diversity would be low, as predicted by the homogeneity assumption. We also expected that this general low interest diversity would still show variance over the range of programs, linked to the motivation of their student populations. Indeed, some programs are chosen through autonomous motivation by students who are highly interested in their program of choice. As students are highly interested in their program of choice, such programs should display a low interest diversity. Other programs could be more attractive to students who have ulterior motives like pleasing their parents and should display a higher interest diversity. As students are less intrinsically interested in their program of choice, these programs display more variance in student vocational interest, resulting in a higher interest diversity. Results for our first question indeed showed that program diversity was low across all study programs, leaving 74% of the higher end diversity continuum unused. On an individual level this find indicates that in general, the PE fit between students and their programs is quite high: RIASEC profiles between student and program correlate .70, on average. Indeed, students predominantly seem to choose a higher education study program that fits their interests quite well when given the opportunity, as is the case in an open access environment. Results also showed that the variance in program interest diversity is related to motivation. Study programs with low interest diversity were linked to students with relatively higher autonomous motivation, while programs with a higher interest diversity were linked to

students with a higher controlled motivation. This relation between student motivation and interest diversity indicates that some programs do attract more students with a higher controlled motivation.

For our second research question, we explored the direct effect of the program interest diversity on average study results. As different programs could reward different interest patterns, we took a conservative approach and pitted three hypotheses against each other on how program interest diversity would influence average program study results. The curvilinear relation between program interest diversity and average study results provided evidence for our mixed effects hypothesis. Different programs indeed rewarded different interest patterns (Smart et al., 2000). To provide an explanation for this curvilinear effect, we performed a post-hoc analysis. Results of this analysis showed that in general, larger program interest diversity was linked to better average study results. In other words, programs with more interest diversity in their student population showed better average results. However, some study programs with very specific interest patterns that scored high on the *Social* dimension and low on the *Realistic* dimension showed an opposite relation: lower program interest diversity in student populations in such environments was associated with better study results. To improve general study results, these findings suggest policy makers and institutions in (open access) higher education should allow for interest diversity in the student population of study programs. At the same time, policy should also ensure a sufficiently high individual student PE interest fit, as literature already suggested (Allen & Robbins, 2010; Assouline & Meir, 1987; Nye et al., 2012; Spokane et al., 2000; Tinsley, 2000; Tracey et al., 2012; Tracey & Robbins, 2006; Tsabari, Tziner, & Meir, 2005). However, to ensure better study results for very specific programs (high on the social dimension and low on the practical dimension) like *Educational Sciences*, the fit between student and program should indeed be as high as possible, resulting in a (very) low program interest diversity and a very

high individual PE interest fit. Finally, results also revealed criterion validity for our continuous approach of program interest diversity: up to 44% of the variance in average study program results can be explained by program interest diversity in student populations. As such, our continuous approach of interest diversity represents a valuable addition to the measurement of internal similarity of environments usually determined through dichotomous test statistics (Bradley-Geist & Landis, 2012; Burke & Dunlap, 2002; Burke et al., 1999).

For our final research question we compared the effect of program interest diversity on study results to the effect of individual PE fit. We hypothesized that due to the low interest program diversity (or high internal similarity) the effect of PE fit would be low. We also tested the cross-level interaction of individual program fit and environmental program diversity. Analyses indicated that the effects at the individual level on study results were very modest at best: student PE fit only explained up to 0.6% at the individual level. Hence, in an open access higher education environment, the variance in PE fit between students and their program barely has a meaningful impact on individual study results. In opposition to these results at the individual level, program interest diversity explained up to 37% of the variance at program level, which is a huge contrast to the observed explained variance at the individual level. Moreover, program interest diversity of different study programs did not only influence average study results, we also obtained partial evidence that this diversity explained more study result variance in the total multilevel models than the individual indicators of PE fit. As only a small part of the total variance in study results (up to 6%) was situated at the program level to begin with, this is no small feat indeed. These findings are analogous to those found in our second question and provide additional evidence that higher education institutions should indeed consider program interest diversity when making policy decisions towards student orientation and admission. As a possible explanation, most students in this open system showed a high PE fit with their program of choice. In systems where choice is restricted

(on the basis of exams or GPA requirements), students may have to choose for programs that match their interests less well, and then this (larger variety in) PE fit has a bigger impact on individual study results. Indeed, earlier research that examined PE fit effects on study results in constrained access systems typically observed more explained variance (Assouline & Meir, 1987; Holland, 1997; Nye, et al., 2012; Spokane, et al., 2000; Tinsley, 2000; Tracey et al., 2012; Tsabari, et al., 2005).

This discussion on the consequences of open access policy illustrates the importance of studying PE fit effects in a variety of study contexts. As Nauta (2010) already indicated, (study) environments remain understudied. Entry exams or GPA requirements yield a selection bias in student intake that will influence the internal similarity in student populations, and therefore also the effects of the observed PE fit variance. Such contextual effects are likely partly responsible for the mixed results regarding the influence of PE fit on study results. For the first time in literature, the present study thus aimed at addressing this problem directly by conducting a PE fit/interest diversity study in a predefined open access environment, firmly rooted in existing theory regarding the possible influence of the environment. As theory predicted, the influence of the environment on outcomes in this open access set up becomes quite influential, while the individual level almost has no explanative power at all regarding study results. In other words, the open access environment causes study program interest diversity to have a profound influence on study results, while severely diminishing the influence of individual student PE fit.

To close the discussion on our third question, the variance in program interest diversity and student PE fit was limited to the extent the cross level interaction between individual and environment was not significant. In other words, program interest diversity did not influence the student PE fit-study results relation: effects of individual PE fit remained low, regardless of the interest diversity of study programs. These findings in our open access

study environment are at odds with results from internal similarity research from Tracey and colleagues (2012). They showed that the effects of PE fit on individual study success was indeed constrained by the study environment. As an explanation, we speculate the open access system leads to such low interest variance that prevents a cross-level interaction between study program interest diversity and student PE fit.

Limitations and Future Research

The present study is unique in its assessment of PE fit effects in an open access system. It would be interesting in the future to directly compare study environments with more constrained entry restrictions on the exact same measurements, using the exact same analyses. We speculate that such an approach would show enlarged PE fit effects in more restricted study programs, while the influence of interest diversity will diminish. The access restrictions could thus be a crucial factor in explaining the mixed findings in literature regarding PE fit, while elaborating literature with interest diversity research.

Our conceptualization of interest diversity and its motivational connection could also be used in organizational and occupational research. We predict that not all work environments will show the same amount of interest diversity. As access to the work environment works quite differently in comparison to access to higher education, we can also expect different effects. For instance, open access to certain jobs (low degree requirements) could result in higher interest diversity. Indeed, a student in an open access environment picks a certain program because that program is of particular interest to him or because his parents wants him to study that specific program. An employee could have other motives to pick a job. Employees who have no (or a low) degree can decide to work out of financial motives exclusively. Though highly speculative, we think that the different motivation in work and study contexts will lead to different patterns of interest diversity for both contexts and could

ultimately end up explaining why the strength of the PE fit – study results relation is so underwhelming in comparison to the theoretical predictions.

Conclusion

In the present study, we have assessed program interest diversity of student populations in study programs. In an open access environment, interest diversity of student populations in study programs is low. RIASEC profiles of students and programs correlate .70 on average. Interestingly, study programs with low interest diversity attract students with relatively higher autonomous motivation, while programs with a higher interest diversity show higher controlled motivation. Despite overall low diversity, the interest diversity of the program environment still had a profound effect on the program's typical study results, while the influence of individual PE fit seems to be nonexistent or very limited at best. In order to enhance student study success, the present study has shown policy makers and educational institutions should focus on a sufficiently high PE fit amongst their student populations, while still allowing for some study program interest diversity.

5.6. References

- 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
- Assouline, M., & Meir, E. I. (1987). Meta-analysis of the Relationship between Congruence and Well-Being Measures. *Journal of Vocational Behavior, 31*(3), 319-332. doi: 10.1016/0001-8791(87)90046-7
- Astin, A. W., Holland, J.L. (1961). The environmental assessment technique – a way to measure college environments. *Journal of Educational Psychology, 52*(6), 308-316. doi:10.1037/h0040137
- Bradley-Geist, J. C., & R. S. Landis (2012). Occupations and Organizations: A Comparison of Alternative Statistical Tests. *Journal of Business and Psychology, 27*(2), 149-159. doi: 10.1007/s10869-011-9233-6
- Brambor, T., Clark, W.R., & Golder, M. (2006). Understanding interaction models: improving empirical analyses. *Political Analysis, 14*(1), 63-82. doi:10.1093/pan/mpi014
- Browne, M. W. (1992). Circumplex models for correlation matrices. *Psychometrika, 57*, 469–497. doi:10.1007/BF02294416
- Burke, M. J., & Dunlap, W. P. (2002). Estimating interrater agreement with the average deviation index: A user's guide. *Organizational Research Methods, 5*(2), 159–172. doi: 10.1177/1094428102005002002
- Burke, M. J., Finkelstein, L. M., & Dusig, M. S. (1999). On average deviation indices for estimating interrater agreement. *Organizational Research Methods, 2*(1), 49–68. doi: 10.1177/109442819921004

- Burns, S. T. (2014). Validity of Person Matching in Vocational Interest Inventories. *The Career Development Quarterly*, 62(2), 114-127. doi: 10.1002/j.2161-0045.2014.00074.x
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. doi: 10.1007/BF02310555
- Donnay, D. (1997). E. K. Strong's legacy and beyond: 70 years of the Strong Interest Inventory. *The Career Development Quarterly*, 46(1), 2-22. doi: 10.1002/j.2161-0045.1997.tb00688.x
- Fonteyne, L., Wille, B., Duyck, W., & De Fruyt, F. (2017). 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*, 17(2), 233-262. doi:10.1007/s10775-016-9327-9
- Graham, S. (2015). Inaugural Editorial for the Journal of Educational Psychology. *Journal of Educational Psychology*, 107(1), 1-2. doi: 10.1037/edu0000007
- Grassi, M., Lucio, R., & Di Blas, L. (2010). CircE: An R implementation of Browne's circular stochastic process model. *Behavior Research Methods*, 42(1), 55-73. doi: 10.3758/BF03192755
- Harris, K. R. (2003). Editorial: Is the work as good as it could be? *Journal of Educational Psychology*, 95, 451-452. doi:10.1037/0022-0663.95.3.451
- Holland, J. L. (1959). A theory of vocational choice. *Journal of Counseling Psychology*, 6, 35-45. doi:10.1037/h0040767
- Holland, J. L. (1966). *The psychology of vocational choice*. Waltham, MA: Blaisdell.
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, FL: Psychological Assessment Resources.

- Hubert, L., & Arabie, P. (1987). Evaluating order hypotheses within proximity matrices. *Psychological Bulletin*, *102*, 172-178.
- Kenny, D. A. (2015). Measuring model fit. Retrieved 21 November 2018, from <http://davidakenny.net/cm/fit.htm>
- King, D. D., Ott-Holland, C. J., Ryan, A. M., Huang, J. L., Wadlington, P. L., & Elizondo, F. (2017). Personality homogeneity in organizations and occupations: considering similarity sources. *Journal of Business Psychology*, *32*(6), 641-653. doi: 10.1007/s10869-016-9459-4
- Linton (1945). *The cultural background of personality*. New York: Century.
- Lounsbury, J. W., Studham, R. S., Steel, R. P., Gibson, L. W., & Drost, A. W. (2009). Holland's Vocational Theory and Personality Traits of Information Technology Professionals. In Y. Dwivedi, B. Lal, M. Williams, S. Schneberger, & M. Wade (Eds.), *Handbook of Research on Contemporary Theoretical Models in Information Systems* (pp. 529-543). Hershey, PA: IGI Global. doi:10.4018/978-1-60566-659-4.ch03
- Lubinski, D., & Humphreys, L. G. (1996). "Seeing the forest from the trees: When predicting the behavior or status of groups, correlate means". *Psychology, Public Policy, and Law*, *2*(2), 363–376. doi:10.1037/1076-8971.2.2.363
- 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

- R Core Team (2018). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL - <http://www.R-project.org/>.
- 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
- Rounds, J. B. (1995). Vocational interests: Evaluation of structural hypotheses. In D. Lubinski & R. V. Dawis (Eds.), *Assessing individual differences in human behavior: New concepts, methods, and findings* (pp. 177–232). Palo Alto, CA: Consulting Psychologists Press.
- Rounds, J., & Su, R. (2014). The Nature and power of Interests. *Current Directions in Psychological Science*, *23*(2), 98-103. doi: 10.1177/0963721414522812
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68-78. doi: 10.1037/0003-066X.55.1.68
- Schneider, B. (1987). The people make the place. *Personnel Psychology*, *40*(3), 437–453. doi: 10.1111/j.1744-6570.1987.tb00609.x.
- Smart, J. C., Feldman, K. A., & Ethington, C. A. (2000). *Academic disciplines: Holland's theory and the study of college students and faculty*. Nashville, TN: Vanderbilt University Press.
- Spokane, A. R., Meir, E. L., & Catalano, M. (2000). Person-Environment Congruence and Holland's Theory: a Review and Reconsideration. *Journal of Vocational Behavior*, *57*(2), 137-187. doi: 10.1006/jvbe.2000.1771

- 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
- Swanson, J. L., & Hansen, J. I. C. (1988). Stability of Vocational Interests over 4-year, 8-year and 12-year Intervals. *Journal of Vocational Behavior*, *33*(2), 185-202. doi:
10.1016/0001-8791(88)90055-3
- 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
- Toomey, K.D., Levinson, E.M., & Palmer, E.J. (2009). A test of Holland's theory of vocational personalities and work environments. *Journal of Employment Counseling*, *46*(2), 82-93. doi:10.1002/j.2161-1920.2009.tb00070.x2
- Tracey, T. J. G., & Robbins, S. B. (2006). The interest-major congruence and college success relation: A longitudinal study. *Journal of Vocational Behavior*, *69*(1), 64-89.
doi:10.1016/j.jvb.2005.11.003
- 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., 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

- Tsabari, O., Tziner, A., & Meir, E. I. (2005). Updated meta-analysis on the relationship between congruence and satisfaction. *Journal of Career Assessment, 13*(2), 216-232.
Doi: 10.1177/1069072704273165
- 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
- Whitney, D. R. (1969), Predicting from expressed vocational choice: a review. *The Personnel and Guidance Journal, 48*(4), 279–286. doi:10.1002/j.2164-4918.1969.tb03318.x

5.7. APPENDIX A:
SIMON-I Questionnaire.

A			
Activities	Dimension	YES	NO
Developing electronic systems	R		
Analysing the grammatical structure of a sentence	I		
Helping people with speech disorders	S		
Recruiting a job candidate	E		
Monitoring the quality standards for food safety and hygiene	C		
Repairing malfunctioning electrical equipment	R		
Carrying out laboratorial analyses	I		
Designing a poster for an exhibition	A		
Helping others with their personal problems	S		
Organising a conference	E		
Preparing financial reports	C		
Being responsible for the maintenance of IT hardware	R		
Analysing statistics	I		
Designing webpages	A		
Developing council prevention campaigns	S		
Presenting new policy propositions	E		
Collecting quantitative and qualitative data	I		
Develop new methods for industrial production	R		
Treating diseases in animals	I		
Editing the sound and images for a movie	A		
Formulating education and training policies	S		
Drawing up the budgets	C		
Doing the follow up on building sites	R		
Analysing x-rays/brain scans	I		
Fit out a show room	A		
Sport guidance for children, the elderly, ...	S		
Formulate a theory about the differences between population groups	I		
Monitor quality standards	C		
Maintaining airplanes	R		
Investigating the impact of historical people	A		
Composing a work of music	A		
Providing guidance for victims	S		
Selling a product or service	E		
Calculating prices	C		
Installing and maintaining computer servers	R		
Designing an advertising folder	A		
Providing information about the assistance for the poor	S		
Drawing up an organisational business or policy plan	E		
Checking bank transactions	C		
Developing windmill parks	R		

Prove a theorem	I
Analysing text structures	A
Giving travel advice	S
Negotiating contracts	E
Drawing up a contract	C
Investigating chromosomal defects	I
Writing scenarios	A
Holding tests, questionnaires and in-depth interviews	S
Screening the administration	C
Working on a drilling rig	R
Turning an idea into a film	A
Giving care to patients	S
Restructuring an organisation or company	E
Checking the compliance of regulations	C
Excluding alternative explanations through experiments	I
Designing the layout of a hospital	A
Advising youngsters regarding their vocational choice	S
Exploring new economic markets	E
Drawing up the annual report	C
Setting up a festival stage	R
Developing a new medicine	I
Writing a review	A
Giving training in communication skills	S
Starting up an enterprise	E
Investigating a cost structure	C
Creating a technical drawing	R
Putting theories in their historical and social context	I
Creating an art piece	A
Giving health advice	S
Giving health and parenting education	E
Calculating expenses	C
Disassembling electrical appliances	R
Comparing cultures	A
Guiding minority groups on the job market	S
Conducting a meeting	E
Drawing up a timetable	C
Measuring a lane	R
Supporting and following up foster families	S
Attracting sponsors	E
Standing in front of a classroom	S
Leading a team	E
Managing a database	C
Collecting soil samples	R
Beginning a herbarium (a plant collection)	I
Counseling underprivileged people	S
Formulating a treatment plan	S

Studying the physical endurance of athletes

I

B

Occupations	Dimension	YES	NO
Industrial designer	R		
Civil engineer	I		
Fashion designer	A		
Policy advisor in political and international relations	E		
Recruitment and selection advisor	E		
Damage expert	C		
Agricultural technician	R		
Teacher	S		
Business economist	C		
Accountant	C		
Electrical engineer	R		
Biologist	I		
Art/music teacher	A		
Speech therapist	S		
Bank manager	C		
Landscape architect	R		
Physicist	I		
Editor	A		
Student counselor	S		
Tax supervisor	C		
Neurologist	I		
Policy advisor art and culture	A		
Educator	S		
Marketing manager	E		
Safety advisor	C		
Construction manager	R		
Historian	I		
Director	A		
Communication manager	E		
Manager (of a company)	E		
Judge	C		
Forester	R		
Researcher	I		
Graphic designer	A		
Psychologist	S		
Lawyer	E		
Notary	C		
Mathematician	I		
Art historian	A		
Social worker	S		
Politician	E		
Pilot	R		
Pharmacist	I		

Linguist	A
Divorce mediator	S
Journalist	A
Structural engineer	R
Lab assistant	I
Photographer	A
Nurse	S
Advertising campaign manager	E
Chemist	I
Tax specialist	C
Architect	R
Artist	A
Educational scientist	S
Librarian	A
Philosopher	I
Representative	E
Geneticist	I
Interior designer	A
Estate agent	E
Physiotherapist	S
Meteorologist	I
Sales manager	E
Statistician	I

Note. Dimensions were masked for the participant and the instructions were as follows. A: 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. B: 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.

Chapter 6. General Discussion.

6.1. General Setting

The process of making the transition towards higher education is a trying endeavor for each student, especially in an open access study environment like Flanders, with nearly unlimited access to all study programs (Tett, Cree, & Christie, 2017). This transition is characterized by two assignments a student has to navigate adequately in order to obtain the desired degree (Fonteyne, 2017). Indeed, making an attainable study choice as well as achieving study success at exams are both crucial towards staying on the model trajectory of the chosen program. These assignments are not straightforward by any means. Reports indicate that about 30% of students in Canada, Europe and the United States leave higher education without getting a degree (OECD, 2017b). Within the current open access higher education structure of this dissertation, I have presented results that only 36% of all first year students manage to stay on track for timely degree attainment (model trajectory + 1 year) after the first year. In contrast, 52% of the students does not stay on track due to failing their exams to various degrees. Only 12% reorients or drops out before taking any exams, presumably because they realize the chosen program does not fit their interests or is unattainable due to an oversized workload in comparison to the already acquired skills in high school.

Ghent University already started the SIMON project in an attempt to improve study success by (re)orienting students at risk to more suitable programs before they fail their exams. Timely reorientation can save a lot of time and resources for both students and institutions for higher education, up to about € 12,000 each year for each student (OECD, 2017a). First results on the (re)orientation advice towards students at risk indicate that (positive and negative) feedback has positive effects on both goal engagement (getting the degree) and disengagement (letting go of unrealistic expectations) in first year students (Fonteyne et al., 2018). The present dissertation aimed at improving the future impact of this SIMON project by addressing a number of issues in literature regarding the two assignments

of study choice and study achievement using the specific focus of a student's study interests. These study interests reflect the vocational interests of students prior to and during the first year of their study program.

6.2. Empirical Findings

In Chapter 2, I have investigated how PE interest correlation fit between a student's study interests and a pool of study programs can be translated into study advice for students that have to make a higher education study choice. This translation is important because literature has indicated that PE (person-environment) interest fit can have a profound effect on study achievement (Tracey & Robbins, 2006; Allen & Robbins, 2010; Nye, Prasad, Bradburn, & Elizondo, 2018; Nye, Su, Rounds, & Drasgow, 2012; Nye, Butt, Bradburn, & Prasad, 2018; Rounds & Su, 2014; Tracey, Allen, & Robbins, 2012). Up until now, there was no objective theoretical or empirical criterion that determined if a study program was interesting enough to advise it to a student. As a result, literature and practice resorted to non-evidence based rules of thumb (Brown & Gore, 1994; Camp & Chartrand, 1992; Tinsley, 2000; Young, Tokar, & Subich, 1998). To address this issue, I have presented an Empirical Advice Set Engine (EASE). EASE is based on the find that PE interest correlation fit (the correlation between the scores on the RIASEC profiles of both student and program, see Tracey et al., 2012 for a more elaborate description) between a student and a pool of study programs approximates a uniform distribution. As such, EASE generates custom made advice sets of study programs for each future student, by balancing the number of study programs with the fit of those study programs. As the results indicate that the algorithm is very stable, the method seemed promising in balancing the size of the advice set against the minimal fit a program has to have in order to be a part of the advice set. About 98% of the future students received an advice set with 7 to 28 options (out of a possible 62), which is still large enough to enable study environment exploration (Holland, 1997). These promising results were confirmed as EASE

was validated by comparing its generated advice sets against advice sets generated by classic methods using high point coding, as used by the SIMON project (Healy & Mourton, 1983; Holland, 1963). EASE generated advice sets of better quality compared to the classic methods. As an example, an advice set of 18 programs delivers study advice to (future) students, explaining about 48% of the variance in the relation between the student's RIASEC profile and the advice set of programs. This level of variance is 16% higher than the level achieved by the classic congruence indices. Besides the direct upgrade in the advice sets for students, EASE also shows strong criterion validity for study programs. Indeed, 81% of successful and persistent students received their own study program as a part of their advice set, which is much higher than similar hit rates reported in literature (Burns, 2014; Donnay, 1997). Besides a high hit rate, it is also important that a method does not generate huge alternative rates or alt rates. An alt rate is a rate that indicates how many times a program is advised to students as a part of their advice set. By enlarging the alt rate, the hit rate will also increase. However, advice sets will also become increasingly larger. As such, balancing the hit rates and the alt rates of programs is crucial to deliver workable advice sets to future students. EASE again outperforms the classic methods balancing both measures. Compared to classic methods, EASE displays a hit rate (at equal alt rates) that is 19% higher at .81, and an alt rate (at equal hit rates) that is 23% lower (which makes the advice sets shorter) at .27. In sum, EASE presents appropriate programs to students, without having to inflate the size of the advice set of study programs.

In Chapter 3, I have investigated a STEM specific study choice, because a recent decline in STEM enrolments is endangering the economic welfare of industrialized regions around the world (Ainly, Kos, & Nicolas, 2008; Perera & McIlveen, 2018). Indeed, STEM higher education does no longer provide a sufficient stream of skilled labor towards the STEM field to meet STEM work place demands (World Economic Forum, 2016). Moreover,

in gender-aware countries, the gender gap is growing further towards an even larger male overrepresentation (Stoet & Geary, 2018; UNESCO, 2016; Xu, 2008). To investigate these issues, STEM research has turned to vocational interests as they are arguably the strongest predictors of study choice (Perera & McIlveen, 2018; Su & Rounds, 2015; Stoll et al., 2017; Su, Rounds, & Armstrong, 2009; Wang & Degol, 2017), alongside other predictors like exposure and performance on high school mathematics (Dejarnette, 2012; Wang, 2013a; Wang 2013b). Remarkably, regarding vocational interests, the brunt of the research has focused on using single RIASEC dimensions and not on PE interest fit of which literature already states it has a profound influence on study choice (Porter & Umbach, 2006). Moreover, as research did not target education specifically, study interests are invaluable to enlarge our knowledge on STEM choice and the gender gap. In Chapter 3, I have thus investigated how study interest fit differs between male and female students. I have also investigated how study interest fit with the STEM (science, technology, engineering, mathematics) field contributes to a STEM study choice and an explanation of the STEM gender gap. Empirical results indicate that female students (vs. male students) have a better fit with their program of choice both inside and outside of the STEM field. A STEM choice prediction model further indicated that the RIASEC dimensions, STEM field interest fit and weekly hours of high school mathematics all showed incremental validity towards a STEM study choice. The gender gap was fully explained through gender interactions on the investigative, social and conventional RIASEC dimensions, weekly hours of high school mathematics and the STEM field interest fit. Most interestingly, a female student's STEM choice is more linked to STEM field fit, while a male student's STEM choice is more linked to weekly hours of mathematics.

In Chapter 4, I have investigated how we can identify students at risk of failing months before they fail their exams in their first year. To this end, I have suggested a novel

methodology that focuses on predicting individual student achievement instead of explaining overall population variance. Indeed, Shmueli (2010) already indicated that the use of both statistical applications (and especially in educational and psychological science) is heavily conflated towards explaining the population variance in study achievement, regardless of the research question. By modeling student failure for each study program specifically, I have also validated a series of predictors for identifying failing students, alongside their already established use in explaining the population variance in study achievement. Predictors included variables that are already effective in the prediction of study achievement (Schneider & Preckel, 2017). For study interests specifically, I have used an adaptation of the polynomial method, complemented by measures of PE interest fit (see Tracey et al., 2012 for an elaboration on Euclidean distance). The program-specific method itself consists of an AIC (Akaike's Information Criterion) procedure that selects the best fitting model from the pool of all models featuring all available predictors (Burnham & Anderson, 2002). This best fitting model is the model that shows the smallest prediction error between the students' actual results and the students' predicted results. A regression of study achievement on the featuring predictors summarizes the selected model. Regarding these models, I have also explored the possibility of different false positive rates. Indeed, as the models aim to identify students at risk of failing in various study programs, it is also important that we limit the identification of passing students as failing. To this extent, I have tested the identification power of different false positive rates. Empirical results show that this program-specific methodology can identify an average of about 29% of the students at risk of failing at a 5% false error rate. Though not the focus of this chapter, the summarizing regression also explained an average of about 37% of the population variance in study achievement. Compared to previous results in SIMON, results explain around two-thirds more population variance (23% vs 37%), while more than doubling the identified number of students at risk of failing from 13% to 29%

(Fonteyne, Duyck, & De Fruyt, 2017). The higher explained variance is probably due to the inclusion of high school GPA (Pinxten, Van Soom, Peeters, De Laet, & Langie, 2017) while it seems reasonable to assume that the higher amount of identified students at risk is (at least partly) a consequence of the prediction-specific methodology. These program-specific prediction models also validate a pool of predictors towards identification of students at risk. Study antecedents (Hodara & Lewis, 2017; Poole, Shulruf, Rudland, & Wilkinson, 2012) and cognitive ability (Rohde & Thompson, 2007; Roth et al., 2015) show the highest potential as they are predictive in nearly all programs. Indeed, these two highly predictive variables form the basis of most prediction models, supplemented with predictors depending on the specific program, like variables of motivation (Kriegbaum, Becker, & Spinath, 2018; Ryan & Deci, 2000; Ryan & Deci, 2017), conscientiousness (De Fruyt & Mervielde, 1996; Furnham, Chamorro-Premuzic, & McDougall, 2003), academic self-efficacy (Bandura, 1993), test-anxiety (Credé & Kuncel, 2008), and to a lesser extent, self-control (Duckworth, Taxer, Eskreis-Winkler, Galla, & Gross, 2019) and metacognition (Kitsantas, Winsler, & Huie, 2008). For study interests specifically, a range of variables including single RIASEC dimensions, correlation fit and Euclidean distance feature in 24% of the program-specific identifying models predicting if students would fail or pass. As such, the collection of study interests form the third most important predictor for identification of failing students after study antecedents and cognitive ability.

In Chapter 5, I have investigated the role of the environment in the PE interest fit relationship and its influence on first year study achievement in an open access environment. Indeed, Nauta (2010) already indicated that the role of the environment was rarely studied, despite its importance to PE interest fit. As such, I have used the homogeneity theory to investigate the study environment of first year students and its influence on study achievement. This theory predicts that students with the same study interests will choose

similar study programs (Schneider, 1987). In this chapter, I have thus found that programs showed a low interest diversity within their student populations. This low interest diversity was due to a generally high average PE interest fit ($r = .70$) of students with their programs of choice. Students indeed seem to choose according to their interests in an open access environment which is completely in line with homogeneity theory (Schneider, 1987). However, between programs, I still observed quite some variance in the interest diversity of programs. I have also found that this variance was tied to autonomous and controlled motivation (Ryan & Deci, 2010; Ryan & Deci, 2017). Programs with a higher diversity were linked to a higher average controlled motivation in students, while programs with a lower diversity were linked to a higher average autonomous motivation. As an explanation, in programs with a higher diversity, students seem to choose less according to their interests and are driven more by exterior motives like pleasing their parents. In programs with lower diversity, students choose more according to their interests, for instance because they actually look forward to studying the chosen subject. This variance in interest program diversity also influenced average program study achievement through a curvilinear relationship. In general, a higher diversity is linked to better achievement, with exception of a particular cluster of programs that score high on the social dimension and low on the realistic dimension. For this cluster, the effect is reversed: a lower diversity is linked to better achievement. To conclude Chapter 5, I have also compared the effects of the environment (program interest diversity) to the effects of the individual (PE interest fit) on study achievement. Surprisingly, I thus found that the effects of individual PE interest fit were nominal at best, while the environment had a profound effect on (average) study achievement.

6.3. Theoretical Impact of the Dissertation

Beside the mere empirical results, this dissertation features a number of properties that have a profound theoretical impact on literature regarding vocational interests. First, all

empirical studies were conducted in an open access study environment (Eurydice, 2017). This open access structure has crucial ramifications for research regarding study interests as the structure of the education system where the research takes place is seldom considered. As an example, access to higher education in Flanders is not locked behind additional requirements like entry exams or GPA requirements. Consequentially, all Flemish students with a high school degree have access to nearly all higher education study programs for a modest yearly tuition fee (Flemish Education in Figures, 2016-2017; OECD, 2017a; OECD, 2017b). As such, the theoretical processes underlying study interests like the homogeneity theory can be studied without unwanted bias from these additional requirements (Holland, 1997; Schneider, 1987). Higher education in Flanders thus presents itself as a unique possibility to investigate study interests under nearly optimal circumstances, which this dissertation has tried to achieve. Results that can be (partly) attributed to this peculiar open access, can also have a profound impact on the literature. For instance, the old adagio of homogeneity theory that students with similar interests will gravitate towards similar environments was nicely illustrated in Chapter 5 (Holland 1997; Schneider, 1987). On average, individual PE interest fit reached a correlation of $r = .70$. This indicates that students do seem to choose according to their study interests if given the chance, which is obviously the case in an open access environment. Moreover, as most students had a good fit with their program, PE interest fit almost did not explain any variance in study achievement, in contrast to the discussed effects at the program level. Although such a finding seems heavily at odds with literature (Tracey & Robbins, 2006; Allen & Robbins, 2010; Nye, Prasad et al., 2018; Nye et al., 2012; Nye, Butt et al., 2018; Rounds & Su, 2014; Tracey et al., 2012), it becomes quite plausible when one takes into account the theories regarding the homogeneity of environments (Holland, 1997; Schneider, 1987). In sum, for future research on study interests, researchers should always take into account the education setup of the environment where the research was conducted.

Second, this dissertation also uses the full RIASEC profile of both students and study programs when investigating study interests, instead of just the dimensions with the highest scores. This full use of all dimensions has always been advocated by Holland himself throughout the development of his RIASEC theory (Holland, 1997; Nauta, 2010). Still, the use of full profiles is not common knowledge in literature and could be (partly) the reason why theoretical expectations regarding the positive effects of PE interest fit on variables like study achievement have remained below expectations (Nye et al., 2012). As a consequence, using all dimensions in this dissertation has had a positive influence on how study interests affect study choice and study achievement. Indeed, Chapter 2 has presented a more accurate and objective way to dispense study advice to future students, clearly trumping the classic coding systems that do not use the full profile (Brown & Gore, 1994; Camp & Chartrand, 1992; Tinsley, 2000; Young et al., 1998). Chapter 3 has presented results that indicate all RIASEC dimensions and their interaction (through Euclidean distance) have an effect on STEM study choice, with four out of six dimensions showing negative effects. Chapter 4 has presented evidence that program-specific prediction models of study achievement tap into a pool of predictors containing all RIASEC dimensions and their interaction(s). In 24% of the programs, one or more RIASEC variables were selected for the final model. The selected variables were not necessarily the highest ones and sometimes even had negative effects on study achievement. For instance, the realistic dimension was not the highest dimension in the psychology program, but did have a (negative) effect on study achievement. Chapter 5 has presented evidence that not only the use of the full student RIASEC profile is important, but also the full RIASEC profile of the environment has to be taken into consideration. As an example, the general find that more program diversity was linked to higher average achievement in study programs was reversed in programs that showed a high social dimension but also a low realistic dimension.

Third, this dissertation not only uses the full RIASEC profile of students and programs to characterize study interests, but also the internal pattern of these dimensions within one profile. Indeed, for the individual student the absolute height of the dimension scores (compared to other students) is less important in comparison to how the RIASEC scores within the student profile relate to each other. This profile pattern is especially crucial regarding orientation towards study choice. Indeed, future students should be advised based on their own RIASEC profile, not on how they score in comparison to others. As such, correlation fit and Euclidean distance both provide a method of making the fit between students and programs, based on internal profile patterns and not necessarily on the absolute height of the RIASEC dimensions (see Chapters 2 and 3). This way, students who have generally lower scores or elevation on all dimensions are not at a disadvantage when receiving study advice.

Fourth, regarding PE correlation interest fit between the study interests of a student and the profile of a program, Chapter 2 has demonstrated that the interest fit between a student's study interests and a set of study programs can be represented by a continuous approximation of a uniform distribution. Literature can now use this property of PE interest fit for further research. To give one example, as the parameters of such an approximation can be estimated, simulation research now becomes possible without the need for actual student and/or program data.

6.4. Practical Impact of the Dissertation

This dissertation on study interests can have a large impact on study orientation applications through the elements of study choice and study achievement. For study choice, the EASE application (Chapter 2) provides a full description on the theory and application of a new and objective criterion which programs should be advised to specific students, based on the PE interest fit between the student's study interests and the program profile. For study

achievement, the program-specific methodology identifies students at risk months before they fail their exams based on the prediction of individual student prediction of achievement (Chapter 4). As a direct consequence, both elements on study choice and study achievement could be practically implemented in future iterations of the SIMON project (Fonteyne, 2017). First, by replacing the classic methods of study choice orientation based on high point coding congruence indices, a number of new features become available. For instance, EASE would facilitate a more fine-grained study choice advice, as the list of selected programs can be further listed from a very good to a very bad PE interest fit. And second, to inform (future) students whether or not a program would be attainable if they decide to study it, the algorithm can now be based the prediction of individual achievement and not on the explanation of population variance.

6.5. Limitations and Future Research

Though this dissertation has a profound impact on both theoretical literature as well as practical application of both study choice and study achievement, the dissertation has one major limitation in addition to some minor ones described in the empirical chapters. It only deals with study achievement of first year students. Although achieving study success is strongly and positively correlated with timely degree attainment (e.g., a high chance of passing the first year corresponds to an 85% chance of obtaining the degree in timely fashion), achieving study success in the first year still remains but a proxy for degree attainment (Fonteyne, 2017). Up until now, the longitudinal data collection within the SIMON project has not yet reached the possibility of making direct predictions regarding degree attainment based on the first stream of data. However, future research on study orientation in general and within SIMON specifically, should focus on predicting achievement towards degree attainment as the degree is the end goal of each student's higher education curriculum. Indeed, partly failing the first year does not necessarily mean the

student will not attain his degree in a timely fashion. In contrast, a student that does pass his first year will not necessarily complete his curriculum.

6.6. Conclusion

In this dissertation, the empirical results and specific operationalization of the RIASEC dimensions provide a unique, open access focus on the effects of study interests. Orientation towards an interesting study choice is now based on an objective criterion of how well a student's study interests have to fit possible programs. Moreover, orientation towards attainable programs is now based on identifying students at risk by focusing on predicting individual study achievement, while still accounting for the specific (open access) set up of the education system. This dissertation also enables counselors to immediately translate these findings towards practical application.

6.7. References

- Ainley, J., Kos, J., & Nicholas, M. (2008). Participation in Science, Mathematics and Technology in Australian Education. *ACER Research Monographs*. Retrieved 11th of June 2019, from http://research.acer.edu.au/acer_monographs/4
- Allen, J. & Robbins, S. (2010). Effects of Interest–Study program Congruence, Motivation, and Academic Performance on Timely Degree Attainment. *Journal of Counseling Psychology*, 57(1), 23-35. doi:10.1037/a0017267
- Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist* 28(2), 117-148. doi:10.1207/s15326985ep2802_3
- Brown, S. D., & Gore, P. A. Jr. (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
- Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A practical information-theoretic approach* (2nd ed.), Springer-Verlag.
- Burns, S. T. (2014). Validity of Person Matching in Vocational Interest Inventories. *The Career Development Quarterly*, 62(2), 114-127. doi: 10.1002/j.2161-0045.2014.00074.x
- Camp, C. C., & Chartrand, J. M. (1992). A comparison and evaluation of interest congruence indices. *Journal of Vocational Behavior*, 41(2), 162–182. doi:10.1016/0001-8791(92)90018-U
- 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

- De Fruyt, F., & Mervielde, Y. (1996). Personality and interests as predictors of educational streaming and achievement. *European Journal of Personality*, *10*(5), 405-425. doi: 10.1002/(SICI)1099-0984(199612)10:5<405::AID-PER255>3.0.CO;2-M
- Dejarnette, N. (2012). America's children: providing early exposure to STEM (Science, Technology, Engineering and Math) initiatives. *Education*, *133*(1), 77–84.
- Donnay, D. (1997). E. K. Strong's legacy and beyond: 70 years of the Strong Interest Inventory. *The Career Development Quarterly*, *46*(1), 2–22. doi: 10.1002/j.2161-0045.1997.tb00688.x
- Duckworth, A. L., Taxer, J. L., Eskreis-Winkler, L., Galla, B. M., & Gross, J. J. (2019). Self-control and Academic Achievement. *Annual Review of Psychology*, *70*, 373-393. doi: 10.1146/annurev-psych-010418-103230
- Eurydice (2017). Retrieved November, 27, 2017, from <https://webgate.ec.europa.eu/fpfis/mwikis/eurydice/index.php/Countries>
- Flemish Education in Figures, 2016-2017, retrieved, June 6th 2019 from <https://www.vlaanderen.be/en/publications/detail/flemish-education-in-figures-2016-2017>
- Fonteyne, L. (2017). *Constructing SIMON : a tool for evaluating personal interests and capacities to choose a post-secondary major that maximally suits the potential*. Ghent University. Faculty of Psychology and Educational Sciences.
- 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. doi: 10.1016/j.lindif.2017.05.003
- Fonteyne, L., Eelbode, A., Lanszweert, I., Roels, E., Schelfhout, S., Duyck, W., & De Fruyt, F. (2018). Career goal engagement following negative feedback : influence of expectancy-value and perceived feedback accuracy. *International Journal of*

Vocational and Educational Guidance, 18(2), 165–180. doi: 10.1007/s10775-017-9353-2

- Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences*, 14(1), 49-66. doi:10.1016/j.lindif.2003.08.002
- Healy, C. C., & Mourton, D. L. (1983). Derivatives of the Self-Directed Search: Potential clinical and evaluative uses. *Journal of Vocational Behavior*, 23, 318-328.
- Hodara, M., & Lewis, K. (2017). How well does high school grade point average predict college performance by student urbanicity and timing of college entry? (REL 2017–250). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest. Retrieved 27th of March 2019 from <http://ies.ed.gov/ncee/edlabs>.
- Holland, J. L. (1963). Explorations of a theory of vocational choice and achievement. II. A four-year prediction study. *Psychological Reports*, 12(2), 547–594. doi:10.2466/pr0.1963.12.2.547
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, FL: Psychological Assessment Resources.
- 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.
- Kriegbaum, Becker, & Spinath (2018). The relative importance of intelligence and motivation as predictors of school achievement: A meta-analysis. *Educational Research Review*, 25, 120-148. doi: 10.1016/j.edurev.2018.10.001

- 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., Butt, S. M., Bradburn, J., & Prasad, J. (2018). Interests as predictors of performance: An omitted and underappreciated variable. *Journal of Vocational Behavior, 108*, 178-189. doi: 10.1016/j.jvb.2018.08.003
- 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
- Nye, C.D., Prasad, J., Bradburn, J., Elizondo, F. (2018). Improving the operationalization of interest congruence using polynomial regression. *Journal of Vocational Behavior, 104*, 154-169. doi: 10.1016/j.jvb.2017.10.012
- OECD (2017a). How Much Do Tertiary Students Pay and What Public Subsidies Do They Receive? Retrieved, October 23rd 2017, from <http://www.oecd.org/edu/>
doi:10.1787/888932461199
- OECD (2017b). The State of Higher Education 2015-2016. Retrieved, June 6th 2019 from <http://www.oecd.org/education/imhe/the-state-of-higher-education-201516.htm>
- Perera, H. N., & McIlveen, P. (2018). Vocational interest profiles: Profile replicability and relations with the STEM major choice and the Big-Five. *Journal of Vocational Behavior, 106*, 84-100. doi: 10.1016/j.jvb.2017.11.012
- Poole P., Shulruf B., Rudland, J., & Wilkinson T. (2012). Comparison of UMAT scores and GPA in prediction of performance in medical school: a national study. *Medical Education 46*(2), 163-171. doi: 10.1111/j.1365-2923.2011.04078.x

- Porter, S.R., & Umbach, P.D. (2006). College Major Choice: An Analysis of Person-Environment Fit. *Research in Higher Education*, 47(4), 429-449. doi: 10.1007/s11162-005-9002-3
- Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, 35, 83-92. doi:10.1016/j.intell.2006.05.004
- Roth, B., Becker, N., Romeyke, S., Schäfer, 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
- Rounds, J., & Su, R. (2014). The Nature and power of Interests. *Current Directions in Psychological Science*, 23(2), 98–103. doi:10.1177/0963721414522812
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. doi: 10.1037/0003-066 x.55.1.68
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. New York: Guilford Publishing.
- Schneider, B. (1997). The people make the place. *Personnel Psychology*, 40(3), 437–453. doi: 10.1111/j.1744-6570.1987.tb00609.x.
- Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, 143(6), 565-600. doi: 10.1037/bul0000098
- Shmueli, G. (2010). *To Explain or to Predict?* *Statistical Science*, 25(3), 289-310. doi: 10.1214/10-STS330
- Stoet, G., & Geary, D. C. (2018). The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education. *Psychological Science*, 29(4), 581-593. doi: 10.1177/0956797617741719

- Stoll, G., Rieger, S., Lüdtke, O., Nagengast, B., Trautwein, U., & Roberts, B.W. (2017). Vocational interests assessed at the end of high school predict life outcomes assessed 10 years later over and above IQ and Big Five personality traits. *Journal of Personality and Social Psychology, 113*(1), 167-184. doi: 10.1037/pspp0000117
- Su, R., & Rounds, J. (2015). All STEM fields are not created equal: People and things interests explain gender disparities across STEM fields. *Frontiers in Psychology, 6*:189. doi: 10.3389/fpsyg.2015.00189
- 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
- Tett, L., Cree, V.E. & Christie, H. (2017). From further to higher education: transition as an on-going process. *Higher Education, 73*(3), 389-406. doi: 10.1007/s10734-016-0101-1
- 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., 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., & Robbins, S.B. (2006). The interest-study program congruence and college success relation: A longitudinal study. *Journal of Vocational Behavior, 69*(1), 64-89. doi:10.1016/j.jvb.2005.11.003
- UNESCO (2016). Sharing Malaysian Experience in Participation of Girls in STEM Education, *Current and Critical Issues in Curriculum and Learning*, retrieved 3rd

October 2018, from <https://unevoc.unesco.org/go.php?q=TVETipedia+Glossary+A-Z&id=640>

Wang, M.T. & Degol, J. L. (2017). Gender Gap in Science, Technology, Engineering, and Mathematics (STEM): Current Knowledge, Implications for Practice, Policy, and Future Directions. *Educational Psychology Review*, 29(1) 119-140.

<https://doi.org/10.1007/s10648-015-9355-x>

Wang, X. (2013a). Modeling Entrance into STEM Fields of Study Among Students Beginning at Community Colleges and Four-Year Institutions. *Research in Higher Education*, 54(6), 664-692. doi: 10.1007/s11162-013-9291-x

Wang, X. (2013b). Why Students Choose STEM Majors: Motivation, High School Learning, and Postsecondary Context of Support. *American Educational Research Journal*, 50(5), 1081-1121. doi: 10.3102/0002831213488622

World Economic Forum (2016). *The Human Capital Report*, retrieved 2 October 2018, from <http://reports.weforum.org/human-capital-report-2016/>

Xu, Y. J. (2008). Gender Disparity in STEM Disciplines: A Study of Faculty Attrition and Turnover Intentions. *Research in Higher Education*, 49(7), 607-624.
doi:10.1007/s11162-008-9097-4

Young, G., Tokar, D. M., & Subich, L. M. (1998). Congruence revisited: Do 11 indices differentially predict job satisfaction and is the relation moderated by person and situation variables? *Journal of Vocational Behavior*, 52(2), 208–223.

doi:10.1006/jvbe.1997.1587

Data Storage Fact Sheets

% Data Storage Fact Sheet 1

% Doctoral dissertation: The Influence of Study Interests and (non-) Cognitive Predictors
on Study Choice and Study Achievement in Flemish Higher Education - Chapter 2

% Author: Stijn Schelfhout

% Date: 20/06/2019

1. Contact details

1a. Main researcher

- name: Stijn Schelfhout

- address: Henri Dunantlaan 2, 9000 Gent

- e-mail: stijn.schelfhout@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Wouter Duyck

- address: Henri Dunantlaan 2, 9000 Gent

- e-mail: wouter.duyck@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:

Stijn Schelfhout, Bart Wille, Lot Fonteyne, Elisabeth Roels, Filip De Fruyt &
Wouter Duyck (2019): From Interest Assessment to Study Orientation: An Empirical Advice Set
Engine, The Journal of Experimental Education, DOI: 10.1080/00220973.2019.1605327

* Which datasets in that publication does this sheet apply to?:

All data

3. Information about the files that have been stored

=====

3a. Raw data

* Have the raw data been stored by the main researcher? YES / NO

If NO, please justify:

* On which platform are the raw data stored?

- researcher PC
- research group file server
- other (specify): external data drive

* Who has direct access to the raw data (i.e., without intervention of another person)?

- main researcher
- responsible ZAP

- 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: ...
- file(s) containing processed data. Specify: spss-files, excel files
- 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
- other: external data drive

* Who has direct access to these other files (i.e., without intervention of another person)?

- main researcher
- responsible ZAP
- all members of the research group
- all members of UGent
- other (specify): ...

4. Reproduction

=====

* Have the results been reproduced independently?: YES / NO

* If yes, by whom (add if multiple):

- name:

- address:

- affiliation:

- e-mail:

% Data Storage Fact Sheet 2

% Doctoral dissertation: The Influence of Study Interests and (non-) Cognitive Predictors
on Study Choice and Study Achievement in Flemish Higher Education - Chapter 3

% Author: Stijn Schelfhout

% Date: 20/06/2019

1. Contact details

1a. Main researcher

- name: Stijn Schelfhout

- address: Henri Dunantlaan 2, 9000 Gent

- e-mail: stijn.schelfhout@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Wouter Duyck

- address: Henri Dunantlaan 2, 9000 Gent

- e-mail: wouter.duyck@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:

Submitted to "Research in Higher Education"

* Which datasets in that publication does this sheet apply to?:

All data

3. Information about the files that have been stored

=====

3a. Raw data

* Have the raw data been stored by the main researcher? YES / NO

If NO, please justify:

* On which platform are the raw data stored?

- researcher PC
- research group file server
- other (specify): external data drive

* Who has direct access to the raw data (i.e., without intervention of another person)?

- main researcher
- responsible ZAP
- 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: ...
- file(s) containing processed data. Specify: spss-files, excel files
- 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
- other: external data drive

* Who has direct access to these other files (i.e., without intervention of another person)?

- main researcher
- responsible ZAP
- 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 Fact Sheet 3

% Doctoral dissertation: The Influence of Study Interests and (non-) Cognitive Predictors
on Study Choice and Study Achievement in Flemish Higher Education - Chapter 4

% Author: Stijn Schelfhout

% Date: 20/06/2019

1. Contact details

1a. Main researcher

- name: Stijn Schelfhout

- address: Henri Dunantlaan 2, 9000 Gent

- e-mail: stijn.schelfhout@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Wouter Duyck

- address: Henri Dunantlaan 2, 9000 Gent

- e-mail: wouter.duyck@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:

Submitted to "Research in Higher Education"

* Which datasets in that publication does this sheet apply to?:

All data

3. Information about the files that have been stored

=====

3a. Raw data

* Have the raw data been stored by the main researcher? YES / NO

If NO, please justify:

* On which platform are the raw data stored?

- researcher PC
- research group file server
- other (specify): external data drive

* Who has direct access to the raw data (i.e., without intervention of another person)?

- main researcher
- responsible ZAP
- 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: ...
- file(s) containing processed data. Specify: spss-files, excel files
- 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
- other: external data drive

* Who has direct access to these other files (i.e., without intervention of another person)?

- main researcher
- responsible ZAP
- 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 Fact Sheet 4

% Doctoral dissertation: The Influence of Study Interests and (non-) Cognitive Predictors
on Study Choice and Study Achievement in Flemish Higher Education - Chapter 5

% Author: Stijn Schelfhout

% Date: 20/06/2019

1. Contact details

1a. Main researcher

- name: Stijn Schelfhout

- address: Henri Dunantlaan 2, 9000 Gent

- e-mail: stijn.schelfhout@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Wouter Duyck

- address: Henri Dunantlaan 2, 9000 Gent

- e-mail: wouter.duyck@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:

Schelfhout S, Wille B, Fonteyne L, Roels

E, De Fruyt F, Duyck W (2019). The effects of vocational interest on study results: Student person – environment fit and program interest diversity.

PLoS ONE 14(4): e0214618. <https://doi.org/10.1371/journal.pone.0214618>

* Which datasets in that publication does this sheet apply to?:

All data

3. Information about the files that have been stored

3a. Raw data

* Have the raw data been stored by the main researcher? YES / NO

If NO, please justify:

* On which platform are the raw data stored?

- researcher PC
- research group file server
- other (specify): external data drive

* Who has direct access to the raw data (i.e., without intervention of another person)?

- main researcher
- responsible ZAP
- 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: ...
- file(s) containing processed data. Specify: spss-files, excel files
- 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
- other: external data drive

* Who has direct access to these other files (i.e., without intervention of another person)?

- main researcher
- responsible ZAP
- all members of the research group
- all members of UGent

- other (specify): ...

4. Reproduction

=====

* Have the results been reproduced independently?: YES / NO

* If yes, by whom (add if multiple):

- name:

- address:

- affiliation:

- e-mail:

Reference List

- ACT, Inc. (2017). The ACT Interest Inventory technical manual. Retrieved October, 23, 2017, from <http://www.act.org/content/dam/act/unsecured/documents/>
- Ainley, J., Kos, J., & Nicholas, M. (2008). Participation in Science, Mathematics and Technology in Australian Education. *ACER Research Monographs*. Retrieved 11th of June 2019, from http://research.acer.edu.au/acer_monographs/4
- Allen, J. & Robbins, S. (2010). Effects of Interest–Study program Congruence, Motivation, and Academic Performance on Timely Degree Attainment. *Journal of Counseling Psychology*, 57(1), 23-35. doi:10.1037/a0017267
- Arbona, C. (2000). Practice and research in career counseling and development – 1999. *The Career Development Quarterly*, 49(2), 98–134. doi:10.1002/j.2161-0045.2000.tb00554.x
- Assouline, M., & Meir, E. I. (1987). Meta-analysis of the Relationship between Congruence and Well-Being Measures. *Journal of Vocational Behavior*, 31(3), 319-332. doi: 10.1016/0001-8791(87)90046-7
- Astin, A. W., & Holland, J. L. (1961). The environmental assessment technique: A way to measure college environments. *Journal of Educational Psychology*, 52(6), 308–316. doi: 10.1037/h0040137
- Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist* 28(2), 117-148. doi:10.1207/s15326985ep2802_3
- Barrick, M. R., & Mount, M. K. (2005). Yes, personality matters: Moving on to more important matters. *Human Performance*, 18(4), 359–372. doi: 10.1207/s15327043hup1804_3

- Blackburn, H. (2017). The Status of Women in STEM in Higher Education: A Review of the Literature 2007–2017. *Science and Technology libraries*, 36:3. doi: 10.1080/0194262X.2017.1371658
- Bradley-Geist, J. C., & R. S. Landis (2012). Occupations and Organizations: A Comparison of Alternative Statistical Tests. *Journal of Business and Psychology*, 27(2), 149-159. doi: 10.1007/s10869-011-9233-6
- Brambor, T., Clark, W.R., & Golder, M. (2006). Understanding interaction models: improving empirical analyses. *Political Analysis*, 14(1), 63-82. doi:10.1093/pan/mpi014
- Brown, S. D., & Gore, P. A. Jr. (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(47), 469-497. doi:10.1007/bf02294416
- Burke, M. J., & Dunlap, W. P. (2002). Estimating interrater agreement with the average deviation index: A user's guide. *Organizational Research Methods*, 5(2), 159–172. doi: 10.1177/1094428102005002002
- Burke, M. J., Finkelstein, L. M., & Dusig, M. S. (1999). On average deviation indices for estimating interrater agreement. *Organizational Research Methods*, 2(1), 49–68. doi: 10.1177/109442819921004
- Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A practical information-theoretic approach* (2nd ed.), Springer-Verlag.

- Burns, S. T. (2014). Validity of Person Matching in Vocational Interest Inventories. *The Career Development Quarterly*, 62(2), 114-127. doi: 10.1002/j.2161-0045.2014.00074.x
- Camp, C. C., & Chartrand, J. M. (1992). A comparison and evaluation of interest congruence indices. *Journal of Vocational Behavior*, 41(2), 162–182. doi:10.1016/0001-8791(92)90018-U
- Cassady, J., & Finch, W. (2015). Using factor mixture modeling to identify dimensions of cognitive test anxiety. *Learning and Individual Differences*, 41, 14 –20. <http://dx.doi.org/10.1016/j.lindif.2015.06.002>
- Cavanaugh, J. E. (1997). Unifying the derivations of the Akaike and corrected Akaike information criteria, *Statistics & Probability Letters*, 31, 201–208. doi:10.1016/s0167-7152(96)00128-9.
- Costa, P.T., & McCrae, R.R. (1992). *Revised NEO Personality Inventory and NEO Five Factor Inventory Professional manual*. Odessa; FL: Psychological Assessment Resources.
- Counseling Psychology*, 45, 204-211. doi:10.1037/0022-0167.45.2.204
- 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
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. doi: 10.1007/BF02310555
- Darcy, M., & Tracey, T.J.G. (2003). Integrating Abilities and Interests in Career Choice: Maximal versus Typical Assessment, *Journal of Career Assessment*, 11(2), 219-237. doi: 10.1177/1069072703011002007

- De Fruyt, F. (2002). A person-centered approach to P–E fit questions using a multiple trait model. *Journal of Vocational Behavior*, *60*(1), 73–90. doi: 10.1006/jvbe.2001.1816
- De Fruyt, F., & Mervielde, Y. (1996). Personality and interests as predictors of educational streaming and achievement. *European Journal of Personality*, *10*(5), 405–425. doi: 10.1002/(SICI)1099-0984(199612)10:5<405::AID-PER255>3.0.CO;2-M
- De Fruyt, F., & Rolland, J. P. (2010). *PfPI. Beschrijving persoonlijkheid op het werk. Handleiding*. Amsterdam/Antwerpen: Pearson.
- Dejarnette, N. (2012). America’s children: providing early exposure to STEM (Science, Technology, Engineering and Math) initiatives. *Education*, *133*(1), 77–84.
- Donnay, D. (1997). E. K. Strong’s legacy and beyond: 70 years of the Strong Interest Inventory. *The Career Development Quarterly*, *46*(1), 2–22. doi: 10.1002/j.2161-0045.1997.tb00688.x
- Duckworth, A. L., Taxer, J. L., Eskreis-Winkler, L., Galla, B. M., & Gross, J. J. (2019). Self-control and Academic Achievement. *Annual Review of Psychology*, *70*, 373–393. doi: 10.1146/annurev-psych-010418-103230
- EUROSTAT. (2017). Annual expenditure on public and private educational institutions per pupil/student. ec.europa.eu/eurostat/tgm/table.do?tab=table&plugin=1&language=en&pcode=tps00067
- Eurydice (2017). Retrieved November, 27, 2017, from <https://webgate.ec.europa.eu/fpfis/mwikis/eurydice/index.php/Countries>
- Flemish Education in Figures, 2016-2017, retrieved, June 6th 2019 from <https://www.vlaanderen.be/en/publications/detail/flemish-education-in-figures-2016-2017>

- Fonteyne, L. (2017). *Constructing SIMON : a tool for evaluating personal interests and capacities to choose a post-secondary major that maximally suits the potential*. Ghent University. Faculty of Psychology and Educational Sciences.
- Fonteyne, L., De Fruyt, F., Dewulf, N., Duyck, W., Erauw, K., Goeminne, K., Lammertyn, J., et al. (2015). Basic mathematics test predicts statistics achievement and overall first year academic success. *European Journal of Psychology of Education, 30*(1), 95–118. doi: 10.1007/s10212-014-0230-9
- 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. doi: 10.1016/j.lindif.2017.05.003
- Fonteyne, L., Eelbode, A., Lanszweert, I., Roels, E., Schelfhout, S., Duyck, W., & De Fruyt, F. (2018). Career goal engagement following negative feedback : influence of expectancy-value and perceived feedback accuracy. *International Journal of Vocational and Educational Guidance, 18*(2), 165–180. doi: 10.1007/s10775-017-9353-2
- Fonteyne, L., Wille, B., Duyck, W., & De Fruyt, F. (2017). Exploring vocational and academic fields of study : development and validation of the Flemish SIMON Interest Inventory (SIMON-I). *International Journal of Vocational and Educational Guidance, 17*(2), 233–262. doi: 10.1007/s10775-016-9327-9
- Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science, 15*(6), 373-378. doi: 10.1111/j.0956-7976.2004.00687.x
- Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences, 14*(1), 49-66. doi:10.1016/j.lindif.2003.08.002

- Gottfredson, G.D., & Holland, J.L. (1975). Vocational choices of men and women – comparison of predictors from self-directed search. *Journal of Counseling Psychology*, 22(1), 28-34. doi:10.1037/h0076150
- Gottfredson, G.D., & Holland, J.L. (1991). *The Position Classification Inventory: Professional manual*, Odessa, FL: Psychological Assessment Resources.
- Gottfredson, L. S. (2002). g: Highly general and highly practical. In R. J. Sternberg, & E. L. Grigorenko (Eds.), *The general factor of intelligence: How general is it?* (pp. 331–380). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Graham, S. (2015). Inaugural Editorial for the Journal of Educational Psychology. *Journal of Educational Psychology*, 107(1), 1-2. doi: 10.1037/edu0000007
- Grassi, M., Luccio, R., & Di Blas, L. (2010). CircE: An R implementation of Browne’s circular stochastic process model. *Behavior Research Methods*, 42(1), 55–73. doi: 10.3758/BRM.42.1.55
- Harris, K. R. (2003). Editorial: Is the work as good as it could be? *Journal of Educational Psychology*, 95, 451-452. doi:10.1037/0022-0663.95.3.451
- Healy, C. C., & Mourton, D. L. (1983). Derivatives of the Self-Directed Search: Potential clinical and evaluative uses. *Journal of Vocational Behavior*, 23, 318-328.
- Hodara, M., & Lewis, K. (2017). How well does high school grade point average predict college performance by student urbanicity and timing of college entry? (REL 2017–250). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest. Retrieved 27th of March 2019 from <http://ies.ed.gov/ncee/edlabs>.
- Holland, J. L. (1959). A theory of vocational choice. *Journal of Counseling Psychology*, 6, 35–45. doi:10.1037/h0040767

- Holland, J. L. (1963). Explorations of a theory of vocational choice and achievement. II. A four-year prediction study. *Psychological Reports, 12*(2), 547–594.
doi:10.2466/pr0.1963.12.2.547
- Holland, J. L. (1966). *The psychology of vocational choice*. Waltham, MA: Blaisdell.
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, FL: Psychological Assessment Resources.
- Hubert, L., & Arabie, P. (1987). Evaluating order hypotheses within proximity matrices. *Psychological Bulletin, 102*, 172-178.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Kenny, D. A. (2015). Measuring model fit. Retrieved 21 November 2018, from <http://davidakenny.net/cm/fit.htm>
- King, D. D., Ott-Holland, C. J., Ryan, A. M., Huang, J. L., Wadlington, P. L., & Elizondo, F. (2017). Personality homogeneity in organizations and occupations: considering similarity sources. *Journal of Business Psychology, 32*(6), 641-653.
doi: 10.1007/s10869-016-9459-4
- 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.
- Kriegbaum, Becker, & Spinath (2018). The relative importance of intelligence and motivation as predictors of school achievement: A meta-analysis. *Educational Research Review, 25*, 120-148. doi: 10.1016/j.edurev.2018.10.001
- Kuder, G. F. (1977). *Activity interests and occupational choice*. Chicago, IL: Science Research Associates.

- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*, 325–343. <http://dx.doi.org/10.3758/s13428-011-0146-0>
- Lent, R. W., Brown, S. D., Hackett, G. (1994). Towards a Unifying Social Cognitive Theory of Academic Interest, Choice and Performance. *Journal of Vocational Behavior*, *45*(1), 79-122. doi: 10.1006/jvbe.1994.1027
- Linton (1945). *The cultural background of personality*. New York: Century.
- Lippa, R. (1998). Gender-related individual differences and the structure of vocational interests: the importance of the people-things dimension. *Journal of Personality and Social Psychology*, *74*(4), 996-1009. doi:10.1037/0022-3514.74.4.996
- Lounsbury, J. W., Studham, R. S., Steel, R. P., Gibson, L. W., & Drost, A. W. (2009). Holland's Vocational Theory and Personality Traits of Information Technology Professionals. In Y. Dwivedi, B. Lal, M. Williams, S. Schneberger, & M. Wade (Eds.), *Handbook of Research on Contemporary Theoretical Models in Information Systems* (pp. 529-543). Hershey, PA: IGI Global. doi:10.4018/978-1-60566-659-4.ch03
- Low, K. S. D., 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), 713–737.
- Lubinski, D., & Humphreys, L. G. (1996). "Seeing the forest from the trees: When predicting the behavior or status of groups, correlate means". *Psychology, Public Policy, and Law*, *2*(2), 363–376. doi:10.1037/1076-8971.2.2.363
- 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

- Noble, J., & Sawyer, R. (2002). Predicting Different Levels of Academic Success in College Using High School GPA and ACT Composite Score. Retrieved 27th of March 2019 from www.researchgate.net/publication/237103886_Predicting_Different_Levels_of_Academic_Success_in_College_Using_High_School_GPA_and_ACT_Composite_Score
- Nye, C. D., Butt, S. M., Bradburn, J., & Prasad, J. (2018). Interests as predictors of performance: An omitted and underappreciated variable. *Journal of Vocational Behavior, 108*, 178-189. doi: 10.1016/j.jvb.2018.08.003
- 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
- Nye, C.D., Prasad, J., Bradburn, J., Elizondo, F. (2018). Improving the operationalization of interest congruence using polynomial regression. *Journal of Vocational Behavior, 104*, 154-169. doi: 10.1016/j.jvb.2017.10.012
- OECD (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. PISA, OECD Publishing, Paris. doi: 10.1787/9789264266490-en.
- OECD (2017). How Much Do Tertiary Students Pay and What Public Subsidies Do They Receive? Retrieved, October 23rd 2017, from <http://www.oecd.org/edu/> doi:10.1787/888932461199
- OECD (2017). Retrieved from <http://www.oecd.org/edu/> doi:10.1787/888932461199
- OECD (2017). The State of Higher Education 2015-2016. Retrieved, June 6th 2019 from <http://www.oecd.org/education/imhe/the-state-of-higher-education-201516.htm>
- 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.

- Päßler, K., & Hell, B. (2012). Do interests and cognitive abilities help explain college major choice equally well for women and men? *Journal of Career Assessment*, 20(4), 479–496. doi: 10.1177/1069072712450009
- Perera, H. N., & McIlveen, P. (2018). Vocational interest profiles: Profile replicability and relations with the STEM major choice and the Big-Five. *Journal of Vocational Behavior*, 106, 84-100. doi: 10.1016/j.jvb.2017.11.012
- Pinxten, M., Van Soom, C., Peeters, C., De Laet, T., & Langie, G. (2017). At-risk at the gate: prediction of study success of first-year science and engineering students in an open-admission university in Flanders—any incremental validity of study strategies? *European Journal of Psychology of Education*, 34(1), 45-66. doi: 10.1007/s10212-017-0361-x
- Poole P., Shulruf B., Rudland, J., & Wilkinson T. (2012). Comparison of UMAT scores and GPA in prediction of performance in medical school: a national study. *Medical Education* 46(2), 163-171. doi: 10.1111/j.1365-2923.2011.04078.x
- Porter, S.R., & Umbach, P.D. (2006). College Major Choice: An Analysis of Person-Environment Fit. *Research in Higher Education*, 47(4), 429-449. doi: 10.1007/s11162-005-9002-3
- Prediger, D. J. (1982). Dimensions underlying Holland’s hexagon: Missing link between interests and occupations? *Journal of Vocational Behavior*, 21(3), 259-287.
- Prediger, D. J. (1998). Is interest profile level relevant to career counseling? *Journal of Counseling Psychology*, 45(2), 204-211. doi:10.1037/0022-0167.45.2.204
- Prediger, D. J. (2000). Holland’s hexagon is alive and well—though somewhat out of shape: Response to Tinsley. *Journal of Vocational Behavior*, 56(2), 197–204.
- R Core Team (2018). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL - <http://www.R-project.org/>.

- Rayman, J., & Atanasoff, L. (1999). Holland's theory and career intervention: The power of the hexagon. *Journal of Vocational Behavior, 55*(1), 114-126.
doi:10.1006/jvbe.1999.1701
- 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
- Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence, 35*, 83-92. doi:10.1016/j.intell.2006.05.004
- Roth, B., Becker, N., Romeyke, S., Schäfer, 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
- Rounds, J. B. (1995). Vocational interests: Evaluation of structural hypotheses. In D. Lubinski & R. V. Dawis (Eds.), *Assessing individual differences in human behavior: New concepts, methods, and findings* (pp. 177–232). Palo Alto, CA: Consulting Psychologists Press.
- Rounds, J., & Su, R. (2014). The Nature and power of Interests. *Current Directions in Psychological Science, 23*(2), 98–103. doi:10.1177/0963721414522812
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78.
doi: 10.1037/0003-066 x.55.1.68
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. New York: Guilford Publishing.
- Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). *High stakes testing in higher education and employment: Appraising the evidence for validity and fairness*. *American Psychologist, 63*(4), 215–227. doi: 10.1037/0003-066X.63.4.215

- Sadler, P. M., Sonnert, G., Hazari, Z., & Tai, R. (2012). Stability and Volatility of STEM Career Interest in High School: A Gender Study. *Science Education*, 96(3), 411–427. doi: 10.1002/sce.21007
- Sawilowsky, S (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, 8(2), 467-474.
- Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., De Fruyt, F., & Duyck, W. (2019a). From Interest Assessment to Study Orientation: an Empirical Advice Set Engine. *The Journal of Experimental Education*. doi: 10.1080/00220973.2019.1605327
- Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., De Fruyt, F., & Duyck, W. (2019b). The effects of vocational interest on study results: Student person – environment fit and program interest diversity. *PLoS ONE*. doi: 10.1371/journal.pone.0214618
- Schneider, B. (1997). The people make the place. *Personnel Psychology*, 40(3), 437–453. doi: 10.1111/j.1744-6570.1987.tb00609.x.
- Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, 143(6), 565-600. doi: 10.1037/bul0000098
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19(4), 460–475. <http://dx.doi.org/10.1006/ceps.1994.1033>
- SDS (2017). Self-Directed Search, a part of Aspira 360. Retrieved October, 23, 2017, from <http://www.self-directed-search.com/>
- Shmueli, G. (2010). *To Explain or to Predict?* *Statistical Science*, 25(3), 289-310. doi: 10.1214/10-STS330
- Smart, J. C., Feldman, K. A., & Ethington, C. A. (2000). *Academic disciplines: Holland's theory and the study of college students and faculty*. Nashville, TN: Vanderbilt University Press.

- Spokane, A.R., Meir, E.L., & Catalano, M. (2000). Person-environment congruence and Holland's theory: a review and reconsideration. *Journal of Vocational Behavior*, 57(2), 137-187. doi: 10.1006/jvbe.2000.177
- Stoet, G., & Geary, D. C. (2018). The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education. *Psychological Science*, 29(4), 581-593. doi: 10.1177/0956797617741719
- Stoll, G., Rieger, S., Lüdtke, O., Nagengast, B., Trautwein, U., & Roberts, B.W. (2017). Vocational interests assessed at the end of high school predict life outcomes assessed 10 years later over and above IQ and Big Five personality traits. *Journal of Personality and Social Psychology*, 113(1), 167-184. doi: 10.1037/pspp0000117
- Su, R., & Rounds, J. (2015). All STEM fields are not created equal: People and things interests explain gender disparities across STEM fields. *Frontiers in Psychology*, 6:189. doi: 10.3389/fpsyg.2015.00189
- 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
- Swanson, J. L., & Hansen, J. I. C. (1988). Stability of Vocational Interest over 4-year, 8-year and 12-year Intervals. *Journal of Vocational Behavior*, 33(2), 185-202.
- Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). "High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72(2), 271–324.
- Tett, L., Cree, V.E. & Christie, H. (2017). From further to higher education: transition as an on-going process. *Higher Education*, 73(3), 389-406. doi: 10.1007/s10734-016-0101-1

- 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
- Toomey, K.D., Levinson, E.M., & Palmer, E.J. (2009). A test of Holland’s theory of vocational personalities and work environments. *Journal of Employment Counseling*, *46*(2), 82-93. doi:10.1002/j.2161-1920.2009.tb00070.x2
- 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. (1995). The Arbitrary Nature of Holland’s RIASEC Types: a Concentric-Circles Structure. *Journal of Counseling Psychology*, *42*(4), 431-439. doi: 10.1037/0022-0167.42.4.431
- 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., 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., & Robbins, S.B. (2006). The interest-study program congruence and college success relation: A longitudinal study. *Journal of Vocational Behavior*, *69*(1), 64-89. doi:10.1016/j.jvb.2005.11.003
- Trapmann, S., Hell, B., Hirn, J.O.W., & Schuler, H. (2007). Meta-Analysis of the Relationship Between the Big Five and Academic Success at University. *Zeitschrift für Psychologie* *215*(2), 132-151. doi:10.1027/0044-3409.215.2.132

- Tsabari, O., Tziner, A., & Meir, E. I. (2005). Updated meta-analysis on the relationship between congruence and satisfaction. *Journal of Career Assessment, 13*(2), 216-232.
Doi: 10.1177/1069072704273165
- UNESCO (2016). Sharing Malaysian Experience in Participation of Girls in STEM Education, *Current and Critical Issues in Curriculum and Learning*, retrieved 3rd October 2018, from <https://unevoc.unesco.org/go.php?q=TVETipedia+Glossary+A-Z&id=640>
- 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
- Vernon, P. A. (1983). Recent Findings on the Nature of g. Vernon, P. A. (1983). *Journal of Special Education, 17*(4), 389–400. doi:10.1177/002246698301700403
- Wang, M.T. & Degol, J. L. (2017). Gender Gap in Science, Technology, Engineering, and Mathematics (STEM): Current Knowledge, Implications for Practice, Policy, and Future Directions. *Educational Psychology Review, 29*(1) 119-140.
<https://doi.org/10.1007/s10648-015-9355-x>
- Wang, X. (2013). Modeling Entrance into STEM Fields of Study Among Students Beginning at Community Colleges and Four-Year Institutions. *Research in Higher Education, 54*(6), 664-692. doi: 10.1007/s11162-013-9291-x
- Wang, X. (2013). Why Students Choose STEM Majors: Motivation, High School Learning, and Postsecondary Context of Support. *American Educational Research Journal, 50*(5), 1081-1121. doi: 10.3102/0002831213488622
- Whitney, D. R. (1969), Predicting from expressed vocational choice: a review. *The Personnel and Guidance Journal, 48*(4), 279–286. doi:10.1002/j.2164-4918.1969.tb03318.x

- Wille, B., De Fruyt, F., & Feys, M. (2010). Vocational interests and Big Five traits as predictors of job instability. *Journal of Vocational Behavior* 76(3), 547-558. doi: 10.1016/j.jvb.2010.01.007
- 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
- World Economic Forum (2016). *The Human Capital Report*, retrieved 2 October 2018, from <http://reports.weforum.org/human-capital-report-2016/>
- Xu, Y. J. (2008). Gender Disparity in STEM Disciplines: A Study of Faculty Attrition and Turnover Intentions. *Research in Higher Education*, 49(7), 607-624.
doi:10.1007/s11162-008-9097-4
- Young, G., Tokar, D. M., & Subich, L. M. (1998). Congruence revisited: Do 11 indices differentially predict job satisfaction and is the relation moderated by person and situation variables? *Journal of Vocational Behavior*, 52(2), 208–223.
doi:10.1006/jvbe.1997.1587