FACULTY OF PSYCHOLOGY
AND EDUCATIONAL SCIENCES

Study orientation in higher education: the Ghent University SIMON project

Mona Bassleer

Supervisor(s)
Prof. Dr. Nicolas Dirix
Prof. Dr. Wouter Duyck
Prof. Dr. Bart Wille

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

Academic year 2023–2024

GHENT
UNIVERSITY

Table of Contents

Voorwoord

Onlangs is mij iets opgevallen. Mijn leeftijdsgenoten kopen een huis of appartement, trouwen, gaan voor gezinsuitbreiding of… lopen een marathon. Ben ik een uitzondering? Ik heb nog geen van deze gebeurtenissen kunnen afvinken. Alhoewel. Een doctoraat schrijven valt misschien toch in het straatje van een marathon lopen, met de juiste dosis verbeeldingskracht.

Toen ik de kans kreeg om te doctoreren, voelde het alsof ik mij inschreef voor een vierjarige marathon met eigenlijk maar weinig hardloopervaring. Toegegeven, ik had reeds intense looptrainingen achter de rug, met onder meer een masterproef als resultaat. Maar was deze voorbereiding voldoende om de lange en uitdagende doctoraatsrace aan te kunnen? Kon ik wel goed genoeg schrijven in academisch Engels? Was ik effectief in staat om de complexe wereld van *big data* te begrijpen en bestuderen? De twijfels waren aanwezig, maar ik voelde ook sterk het geloof van anderen in mijn capaciteiten en werd herinnerd aan één van de quotes uit mijn notities: “Ik wil geen spijt hebben van zaken die ik niet heb geprobeerd.” Dus zo zette ik uiteindelijk toch de stap richting de startlijn.

Het eerste jaar moest ik op gang komen. Ik liep wel, maar achteraf gezien toch wat traag en in een mistig landschap, niet goed wetende waar naartoe precies. Het was ook vreemd om een dergelijke marathon te starten in de schaduw van een wereldwijde pandemie.

Mijn verwachting om deel uit te gaan maken van een bruisende academische gemeenschap moest snel worden getemperd. Het parcours lag er wat verlaten bij.

Tijdens de twee volgende jaren begon de mist op te trekken en kreeg ik meer ritme (lees: structuur). De aanwezigheid van medelopers werd duidelijker en hun energie werkte aanstekelijk. Maar bij een marathon horen ook obstakels. Een gezondheidsissue voelde als een onverwachte blessure tijdens mijn race. Ik moest af en toe eens wandelen om op adem te kunnen komen. Desondanks bleef ik de verfrissende waterposten wel halen. Hernieuwde kracht dook telkens op wanneer een analyse lukte of een paper beter vorm kreeg.

Het laatste jaar bracht de finishlijn dichterbij. Maar bestond deze weg nu niet uit iets steilere heuvels? Of maakten de blokkades die sommige lopers (editors en reviewers) opwierpen het mij gewoon soms te lastig? Stress en druk namen ongewild meer en meer de bovenhand. Op een bepaald moment raakte ik toch wel even de ‘beruchte muur’. Ik botste, maar kaatste terug. Ik werd opgevangen en terug op mijn benen gezet. Dankzij mijn ondersteuningsteam. Zij verdienen een bijzonder woord van dank.

Allereerst bedank ik graag mijn (co)promotoren *Wouter Duyck* en *Nicolas Dirix*, en niet-officiële copromotor *Stijn Schelfhout*. Jullie waren de coaches die mij de weg wezen, begeleidden en bleven aanmoedigen tijdens mijn marathon.

Wouter, bedankt om mij vier jaar geleden deze marathonopportuniteit te geven en mij deel uit te laten maken van jouw lab. Ik ben zo onder de indruk van jouw kennis en ervaring en mag van geluk spreken dat je deze met mij wilde delen. Als ik heel eerlijk ben, zag ik jou in het begin misschien als een strenge coach. Maar eigenlijk wilde ik jou gewoon absoluut niet teleurstellen. Ik was tenslotte geen psycholoog, maar de eerste pedagoog in jouw team. Wat een eer. Naarmate ik meer kilometers aflegde, voelde ik echt jouw appreciatie voor mijn prestaties. Dit deed deugd. Je reikte mij ideeën en concrete handvaten aan wanneer ik hier nood aan had tijdens mijn race. Ideeën en handvaten die nooit werden opgedrongen, maar vergezeld werden met oprechte interesse in mijn mening en argumentatie.

Nicolas, na het startschot nam jij vrij snel de ‘hoofdcoachfakkel’ over van Wouter. En hoe. Bedankt voor jouw open houding. Ongeacht hoe triviaal mijn vragen en twijfels soms leken, jij creëerde een veilige en ondersteunende omgeving waarin ik op mijn gemak was. Waarin ik mij de race zag uitlopen. Gekarakteriseerd door jouw luisterend oor en bereidheid om steeds te zoeken naar mogelijke strategieën en oplossingen. Zeker wanneer de heuvels voor mij steiler aanvoelden. En ook op momenten die voor jou misschien minder goed uitkwamen. Bedankt om regelmatig te vragen hoe het met mij ging tijdens het lopen. Aandacht te hebben voor mijn ‘paniekskes’. Om mij mijn ritme te laten terugvinden wanneer ik deze even kwijt was. En mij erop te wijzen dat af en toe wandelen meer dan oké en zelfs nodig is.

Stijn, ik introduceerde jou reeds als mijn niet-officiële copromotor. Misschien dan niet geregistreerd als mijn coach, maar daarom niet minder van belang geweest. Integendeel. Zeker gezien jij een aantal jaren geleden een race voor hetzelfde project hebt gelopen. En dus als geen ander wist hoe deze marathon kon aanvoelen. Wanneer ik geconfronteerd werd met blokkades, stond jij steeds als eerste in de rij om een peptalk vol woorden van her- en erkenning te geven. Juist wat ik op dergelijke momenten nodig had. Om verder te kunnen lopen richting de eindstreep. Dat heb ik enorm geapprecieerd. Ook bedankt om mij onderweg van inzichten en entertainment te voorzien. Als een echte taalvirtuoos en datagoeroe.

Verder wil ik mijn copromotor *Bart Wille*, *Filip De Fruyt* als lid van mijn begeleidingscommissie, *Eleonore Smalle* als lid van het lab, en *Ilse De Bourdeaudhuij*, *Maya Caen*, *Mit Leuridan* en *Lot Fonteyne* als co-auteurs bedanken voor de feedback en bemoedigende woorden die ze mij vanaf de zijlijn gaven. Om zo optimaal mogelijk mijn einddoel te kunnen halen. Eveneens een welgemeende dankjewel aan *Lies* en *Christophe* voor de administratieve en technische ondersteuning hierbij. *Lot*, jij ook bedankt om destijds het SIMON-project samen met Wouter leven in te blazen. En mijn deelname aan deze marathon dus mogelijk te maken.

Daarnaast kan ik natuurlijk mijn vaste medelopers niet vergeten. *Sofie* en *Merel*, wat prijs ik mijzelf gelukkig dat jullie mee met mij aan de startlijn stonden. En heel de tijd in mijn buurt zijn gebleven. En *Cathy*, dat jij ons na het eerste jaar hebt vergezeld. Bedankt voor de vele deugddoende gesprekken en duwtjes in de rug. Om na de zoveelste keer toch nog eens mijn frustraties en irritaties te aanhoren. Jullie zorgden ervoor dat ik tussen het, soms misschien wel iets te hard, focussen, mijn lachspieren niet vergat. Ik wens iedereen medelopers zoals jullie toe. En wacht jullie met de grootste trots op aan de eindmeet van jullie marathonraces. *Samuël*, *Alexandra*, *Sam*, *Pieter*, *Liam* en *Julie*, ook jullie bedankt om bij momenten naast mij te komen lopen en mij te verrijken met jullie empathie en verhalen.

Eveneens een speciale bedanking aan mijn vriendengroep, en in het bijzonder *Silke*, *Elina* en *Goedele*, en (schoon)familie die deel uitmaakten van mijn fantastische supportersclub. Bedankt om af en toe met mij een pauze te willen nemen. Pauzes die onmisbaar waren en zeer verkwikkend werkten.

Nu ik richting het einde van mijn dankbetuigingen ga, kan een dankwoord voor mijn trouwste supporters uiteraard niet ontbreken. De crew van mijn volgwagen.

Mama en *papa*, bedankt om mij al gedurende heel mijn leven te steunen in mijn beslissingen. Om mijn bakermat te zijn. Zodat ik deze marathon kon en durfde ontdekken. Ook dankjewel *zus* om samen met mama en papa plaats te nemen in mijn volgwagen. En samen ervoor te zorgen dat ik tussendoor voldoende hydrateerde en rustmomenten inlaste.

Tot slot. Van onschatbare waarde. Lieve *Mathias*. Cliché verwoording, maar waar moet ik beginnen? Jij bezit een buitengewone gave. Jij fungeerde als één van mijn coaches. Maar ook als leider van mijn supportersclub langs de zijlijn. Als bestuurder van mijn volgwagen. En je liep ook nog eens met mij mee. Terwijl je zelf ook een andere marathon loopt. Of twee. Of drie. Bij kleine overwinningen tijdens mijn race was jouw enthousiasme soms groter dan de mijne. Wanneer ik mijn evenwicht verloor, ving jij mij als eerste op. En hielp je mij opnieuw de juiste richting vinden. Bedankt om mijn relativeringsvermogen verder aan te wakkeren. De obstakels en blokkades op mijn weg waren inderdaad leermomenten. Enkel in vermomming. Dus geen onoverkomelijke muren. En ik ben er inderdaad geraakt. Aan de eindmeet. Een vierjarige marathon uitgelopen. Dan zou ik zeggen… tijd om één of meerdere van die andere gebeurtenissen ook te gaan afvinken?

Gent, 20 juni 2024

Mona Bassleer

Samenvatting

Het hoger onderwijssysteem in Vlaanderen wordt gekenmerkt door zijn relatief open toegankelijkheid en betaalbaarheid. Echter, de overvloed aan studieopties die dit systeem met zich meebrengt, kan het besluitvormingsproces van studenten met betrekking tot hun studiekeuze richting hoger onderwijs compliceren. Deze complexiteit kan op zijn beurt de succespercentages in het hoger onderwijs beïnvloeden en kosten inhouden voor studenten, gezinnen en de samenleving. Daarom ontwikkelde Universiteit Gent de studie(her)oriënterings- en remediëringstool SIMON (Studiecapaciteiten en Interesse MONitor), die in 2012 werd gelanceerd en vanaf 2015-2016 breed werd ingezet. Dit online zelfbeoordelingsplatform helpt toekomstige studenten hoger onderwijs bij het kiezen van een studie die past bij hun interesses en (niet-)cognitieve competenties zoals zelfcontrole en wiskundige bekwaamheid, terwijl SIMON ook gepersonaliseerde begeleiding biedt aan nieuw ingeschreven studenten in het hoger onderwijs op basis van hun competenties. De tool helpt zo bij het identificeren en tijdig ondersteunen van studenten met academische vaardigheidstekorten. Sinds de oprichting heeft SIMON reeds gegevens verzameld van meer dan 70,000 studenten, wat voortdurende verbeteringen mogelijk maakt en zijn relevantie voor de evoluerende onderzoeks- en onderwijslandschappen te waarborgen. Bijgevolg maakte het huidige proefschrift gebruik van (recente) inzichten uit theorie en methodologie samen met deze uitgebreide SIMON-gegevens om de componenten van de tool voor het beoordelen van

interesses (SIMON-I) en competenties (SIMON-C), en zijn feedbackmodule te optimaliseren. Hiermee streefden we ernaar bij te dragen aan het doel van SIMON om de academische trajecten van (toekomstige) studenten te verbeteren, ten gunste van onderzoek en onderwijspraktijk.

Aanvankelijk concentreerden we ons in Hoofdstuk 2 op de verfijning van SIMON-C door de effecten van de COVID-19-pandemie op academische prestaties in het hoger onderwijs te onderzoeken, inclusief mogelijke interacties tussen de pandemie en achtergrond/(niet-)cognitieve studentenkenmerken, bekend als determinanten van academische prestaties. We beschouwden SIMON-C-data van eerstejaarsstudenten uit de vier pre-pandemische cohorten 2015-2016 tot en met 2018-2019, cohort 2019-2020 dat één derde van hun eerste academiejaar tijdens de pandemie doorbracht en cohort 2020-2021 dat het volledige eerste academiejaar tijdens de pandemie ervaarde. De bevindingen tonen aan dat studenten uit cohort 2019-2020 hogere academische prestaties laten zien in vergelijking met de pre-pandemische cohorten. Daarentegen toont cohort 2020-2021 de laagste academische prestaties in vergelijking met drie van de vier pre-pandemische cohorten en cohort 2019-2020. Socio-economische status (SES) modereert ook het effect van de pandemie op academische prestaties, waarbij de grootste kloof in academische prestaties tussen studenten met een lage SES en hoge SES wordt waargenomen bij cohort 2020-2021.

In Hoofdstuk 3 hebben we de Ghent University Language Screening (GULS) voor SIMON-C geïntroduceerd en gevalideerd. GULS is een open toegankelijke Nederlandse taalbeoordelingstest die specifiek is ontworpen om de academische taalvaardigheid in termen van leesbegrip onder eerstejaarsstudenten in het hoger onderwijs te evalueren na inschrijving. De 18-item GULS toont sterke constructvaliditeit en betrouwbaarheid, en is in het bijzonder effectief bij het beoordelen van studenten met lagere academische taalvaardigheid. Bovendien draagt GULS bescheiden bij aan de voorspelling van academische prestaties overheen verschillende basis en meer gevorderde wiskundige studieprogramma's en overheen beide categorieën afzonderlijk, ook bovenop achtergrond en andere cognitieve studentenkenmerken. Afhankelijk van het studieprogramma varieert de bijdrage van GULS aan academische prestaties. Conform de verwachtingen, manifesteert het studieprogramma 'Toegepaste Taalkunde' zich duidelijk. Enkel voor 'Biochemistry and Biotechnology' and 'Engineering Technology', speelt GULS geen rol in de predictie van academische prestaties.

Verder, in Hoofdstuk 4, benadrukten we de feedbackmodule geassocieerd met SIMON-C. Onze aandacht was gericht op de betrokkenheid van eerstejaarsstudenten bij de ontvangen feedback over hun voorspelde slaagkans in het eerste jaar, aangevuld met aanbevelingen voor remediërings/competentietrainingsactiviteiten, en de onderliggende mechanismen van deze betrokkenheid. We testten een model gebaseerd op de Theorie van Gepland Gedrag *(Theory of Planned Behavior)*, waarbij de ontvangen feedback van studenten, hun feedbackzelfeffectiviteit (vertrouwen in eigen capaciteiten om met feedback om te gaan), en zowel hun intentie om met de feedback om te gaan als hun daadwerkelijke feedbackbetrokkenheidsgedrag werden meegenomen. Onze bevindingen tonen aan dat studenten die van plan zijn zich met de ontvangen feedback bezig te houden, dit ook waarschijnlijker zullen doen. Bovendien beïnvloedt feedbackzelfeffectiviteit de intentionele en gedragsmatige feedbackbetrokkenheid van studenten, maar deze laatste impact is volledig afhankelijk van de aanwezigheid van hun gerelateerde intenties. Ook tonen studenten die feedback ontvingen met een (vrij) hoge slaagkans hogere feedbackzelfeffectiviteit, wat op zijn beurt nodig blijkt om intentioneel met de feedback om te gaan.

Ten slotte hebben we ons in Hoofdstuk 5 gefocust op SIMON-I. De logistisch geregresseerde persoon-omgevingsinteressefit (*logistic regressed person-environment (PE) interest fit (LRIF)*) werd gepresenteerd, als een methode voor het beoordelen van de overeenkomst tussen de interesseprofielen van een student en een studieprogramma. Specifiek beschouwt LRIF zowel de interessepatronen die overeenkomen met een specifieke omgeving als de intressepatronen die hiervan afwijken. Daarnaast vereist LRIF geen aanvullende gegevensverzameling om omgevingsinteresseprofielen te bepalen. We hebben onderzocht hoe goed LRIF studiekeuze voorspelt, ook in vergelijking met meer traditionele methoden met betrekking tot persoon-omgevingsinteressefit zoals *Euclidean distance PE interest fit (EDF)* en *correlation PE interest fit (CF)*. De resultaten geven aan dat LRIF nauwkeurig onderscheidt tussen studenten die een bepaald studieprogramma kiezen en diegenen die een ander studieprogramma kiezen. Bovendien is LRIF gelijkwaardig aan CF en presteert zelfs beter dan EDF bij het voorspellen van studiekeuze.

Ter conclusie draagt dit proefschrift bij aan onderzoek en onderwijspraktijk op meerdere domeinen en biedt, zoals we zouden kunnen zeggen, 'begeleiding voor betere studiebegeleiding’ richting/in hoger onderwijs. Pandemie-gerelateerde leerachterstanden in het hoger onderwijs zijn gering, maar het erkennen van de kwetsbaarheid onder studenten met een lage SES en hun ondersteunen blijft belangrijk. GULS kan worden ingezet als tool voor toekomstig populatieonderzoek en studie(her)oriënterings- en remediëringsadvies met betrekking tot Nederlandse taalvaardigheid. Educatieve interventies kunnen zich richten op het versterken van de feedbackzelfeffectiviteit van studenten om op die manier hun feedbackbetrokkenheid te verbeteren. De LRIF-methode biedt een alternatief voor het meten van persoon-omgevingsinteressefit dat minstens zo voorspellend is voor studiekeuze als meer traditionele methoden, maar geavanceerder en tijdsefficiënter voor toepassing in studie- (en loopbaan)keuzebegeleidingsomgevingen zoals SIMON.

Summary

Flanders' higher education system is known for its relative open accessibility and affordability. However, the plethora of study options available under this system can complicate the study choice decision-making process towards higher education for students. This complexity can, in turn, affect higher education success rates and imposing costs on students, families, and society. Ghent University therefore developed the study (re)orientation and remediation tool, SIMON (Study capacities and Interest MONitor), launched in 2012 and widely deployed from 2015-2016. This online self-assessment platform assists prospective students in choosing a study that suits their interests and (non-)cognitive competences like self-control and mathematical proficiency, while also offering personalized guidance to newly enrolled higher education students based on their competences. As such, the tool helps identify and timely support students with academic skill deficiencies. Since its inception, SIMON has collected data from over 70,000 students, allowing continual improvements and maintain its relevance to the evolving research and education landscapes. Therefore, the present dissertation used (recent) insights from theory and methodology along with this extensive SIMON data to optimize the tool's components for assessing interests (SIMON-I) and competences (SIMON-C), and its feedback module. In doing so, we aimed to contribute to SIMON's goal of improving (prospective) students' academic trajectories, thereby benefiting research and educational practice.

Initially, in Chapter 2, we focused on the refinement of SIMON-C by investigating the impacts of the COVID-19 pandemic on academic achievement in higher education, including possible interactions between the pandemic and student background/(non-)cognitive characteristics, known as determinants of academic achievement. We considered SIMON-C data of first-year university students from the four pre-pandemic cohorts 2015-2016 to 2018-2019, cohort 2019-2020 who spent one-third of their first academic year during the pandemic, and cohort 2020-2021 who experienced the entire first academic year during the pandemic. The findings demonstrate that students from the 2019-2020 cohort demonstrate higher academic achievement compared with the pre-pandemic cohorts. Conversely, the 2020-2021 cohort shows the lowest academic achievement when compared with three of the four pre-pandemic cohorts and the 2019-2020 cohort. Socio-economic status (SES) is also found to moderate the pandemic's effect on academic achievement, with the largest academic achievement gap between low and high SES students for the 2020-2021 cohort.

In Chapter 3, we introduced and validated the Ghent University Language Screening (GULS) for SIMON-C. GULS is an open access Dutch post-entry language assessment test specifically designed to evaluate academic language proficiency in terms of reading comprehension among first-year higher education students. The 18-item GULS shows strong construct validity and reliability, particularly effective at assessing students with lower academic language proficiency. In addition, GULS modestly contributes to the prediction of academic achievement across various basic and more advanced mathematics study programs and across both categories separately, also beyond background and other cognitive student characteristics. Depending on the field of study, GULS's contribution to academic achievement varies. Not surprisingly, the study program Applied Language Studies stands out. Only for Biochemistry and Biotechnology and Engineering Technology, GULS plays no role in the prediction of academic achievement.

Further, in Chapter 4, we emphasized the feedback module associated with SIMON-C. Our attention was directed towards the engagement of first-year university students with the received feedback on their predicted first-year chance of study success, supplemented by recommendations for remedial/competence training activities, and its underlying mechanisms. Therefore, we tested a model based on the Theory of Planned Behavior, including students' received feedback, feedback self-efficacy (i.e., confidence in one's own abilities to engage with feedback), and both their intention to engage with the feedback as well as their actual feedback engagement behavior. Our findings show that students who plan to engage with the received feedback are more likely to actually do so. Additionally, feedback self-efficacy influences students' intentional and behavioral feedback engagement but the latter impact fully relies on the presence of their related intentions. Also, students who received feedback with a (fairly) high chance of study success exhibit higher feedback self-efficacy, which in turn is needed to engage with the feedback intentionally.

Finally, in Chapter 5, we concentrated on SIMON-I and presented logistic regressed person-environment (PE) interest fit (LRIF) as a method for assessing the match between the interest profiles of a student and a study program. Specifically, LRIF considers the interest patterns that align with but also those that diverge from a specific environment, and does not require additional data collection to establish environment interest profiles. We investigated how well LRIF predicts study choice, also compared with more traditional PE interest fit methods such as Euclidean distance PE interest fit (EDF) and correlation PE interest fit (CF). The results indicate that LRIF accurately distinguishes between students who choose a particular study program and those who choose another one. Furthermore, LRIF equals CF and even outperforms EDF in predicting study choice.

To conclude, the present dissertation contributes to research and educational practice in multiple areas and could be said to provide 'guidance for better study guidance' towards/in

higher education. Pandemic-related learning losses in higher education are minor but recognizing the vulnerability among low SES students and supporting them remains important. GULS can be used as a tool for future population-level research and study (re)orientation and remediation advice regarding Dutch language proficiency. Educational interventions can focus on boosting students' feedback self-efficacy to enhance their feedback engagement. The LRIF method offers an alternative to measure PE interest fit that is at least as predictive for study choice as more traditional methods, yet more advanced and time-efficient for application in study (and career) choice counseling settings like SIMON.

1

Introduction

The transition from secondary to higher education is far from easy for (prospective) students (Tett et al., 2017; van Daal et al., 2013). When enrolling in Flemish higher education specifically, students encounter a relative open access system (Eurydice, 2023; OECD, 2021a) with low enrollment fees (Kelchtermans & Verboven, 2010; OECD, 2021a), typically less than €1,000 per year (Study in Flanders, 2024). This structure allows individuals holding a secondary education diploma to enter a higher education study program, with the exceptions of Medicine, Dentistry, and Performing and Visual Arts, which require entrance exams[1] (Vlaamse overheid, 2024). On the one hand, this system presumes to ensure equitable access to higher education, fostering the involvement of economically disadvantaged demographics. Students are also offered a wide array of higher education opportunities, regardless of their secondary education specialization. On the other hand, however, the limited formal admission requirements place the responsibility for study choice on the student (Fonteyne, 2017, 2022). Additionally, the extensive range of choices can trigger an overwhelming and experimenting decision-making process among students (Oppedisano, 2009) and contribute to high dropout and low success rates (Fonteyne, 2017; Schelfhout, 2019). Consequently, these outcomes carry implications for students, parents, and society in terms of resources, effort, and time (Fonteyne, 2017; OECD, 2022; Schelfhout, 2019). For instance, in 2019, a Belgian higher educational institution invested $21,081.70 per student (OECD, 2022).

To improve students' academic trajectories, adequate support in their challenging process of study (re)orientation and remediation towards/in higher education is recommended (Fonteyne, 2017; Schelfhout, 2019). Alongside Flemish government initiatives like Columbus, researchers at Ghent University[2] therefore developed the non-binding, free, and online self-assessment tool SIMON (Study capacities and Interest MONitor). SIMON can be

[1] As of the academic year 2023-2024, an entrance exam is also mandatory for Veterinary Medicine (Vlaamse overheid, 2024).
[2] In collaboration with HOGENT, Howest and Arteveldehogeschool.

summarized as a (re)orientation and remediation tool that aims to help prospective students (i.e., last-year secondary education students) making a study choice that aligns with their interests and (non-)cognitive competences (e.g., test anxiety and mathematical proficiency), and to offer targeted guidance to first-year higher education students based on their (non-)cognitive competences (Fonteyne, 2017, 2022). As such, particularly students at risk of having insufficient basic skills for academic success can be identified and supported in a timely manner (Fonteyne, 2017; Schelfhout, 2019). SIMON originated in 2012 and is implemented on a large-scale from 2015-2016 onwards (Fonteyne, 2022). Prior research reveals that SIMON manages to accurately identify 29% of the at-risk students (at a 5% false positive rate[3]) (Schelfhout, 2019).

Meanwhile, SIMON has gathered a unique and extensive historical dataset of more than 70,000 students regarding study choice, student characteristics (background, (non-)cognitive, interests), and academic achievement metrics (Fonteyne, 2022). Additionally, ongoing literature offers new insights and suggestions for future research on academic achievement, feedback engagement and study choice, while society also continues to develop and encounters unforeseen situations (e.g., COVID-19 pandemic). Flemish higher education statistics further demonstrate that in 2022-2023, first-year students in both professional and academic bachelor's programs earn an average of 63% of their enrolled ECTS[4] credits, reflecting a trend observed since 2009-2010 (Statistiek Vlaanderen, 2024b). However, the percentage of students who graduate within the expected timeframe of three years decreases from 36% among the students who started their bachelor's program in 2008-2009 to 30% among those who started in 2020-2021 (Statistiek Vlaanderen, 2024a). Across

[3] Students that are successful but identified as failing.

[4] European Credit Transfer and Accumulation System. ECTS facilitates the transfer of credits between higher education institutions, enabling credits earned at one institution to count towards a qualification pursued at another. These credits reflect learning outcomes and associated workload (European Commission, 2021).

OECD countries, this percentage stands at 39% for full-time students who enrolled in their bachelor's program in 2018-2019 (OECD, 2022).

Therefore, continuously studying and optimizing SIMON is crucial to ensure its effectiveness in addressing the (evolving) needs of and insights from research, the educational landscape, and its actors. By updating and refining features based on user feedback and conceptual and methodological advancements, SIMON can better support students in making informed decisions about their academic paths and provide more fine-tuned targeted guidance at the start of higher education. Such optimization thus maintains relevance, accuracy, and usability, ultimately enhancing the overall experience and outcomes for students using the tool.

The present dissertation applies (newly) acquired conceptual and methodological insights from literature to SIMON data to optimize the tool and so further improve students' academic journeys, thereby bridging gaps in scientific research and contributing to educational practice. Four empirical chapters (Chapter 2 to Chapter 5) are presented, alongside this general introduction (Chapter 1) and the general discussion (Chapter 6). The empirical chapters each concentrate on a specific aspect of SIMON, aiming to comprehensively optimize its functionality and accuracy for the benefit of the student (and society). Specifically, Chapter 2 and Chapter 3 focus on the competence component (SIMON-C), Chapter 4 on the corresponding feedback module, and Chapter 5 on the interest component (SIMON-I). In short, Chapter 2 addresses the case of the unforeseen COVID-19 pandemic and its effects on first-year academic achievement in higher education. Chapter 3 contributes to the validation of the Ghent University Language Screening (GULS) that assesses Dutch language proficiency among first-year academic achievement to identify and support at-risk students. Chapter 4 focuses on the examination of determinants and underlying mechanisms of student feedback engagement based on a Theory of Planned Behavior model.

Finally, Chapter 5 presents a method to fit the vocational interest profiles of students and study programs that encompasses both interest patterns that align with and divert from a specific environment, and eliminates the need for additional data collection to obtain environment interest profiles.

The SIMON Project

The SIMON project encompasses three applications. The first application, ‘Vraag het aan SIMON’, assists secondary education students in their study choice process towards higher education. These prospective students receive answers to two key questions: "Which higher education study programs match my interests?" and "What are my chances of success in these study programs?". The second application, ‘SIMON zegt’, specifically supports first-year students at Ghent University. Upon completion, these students are provided with a personal feedback report, detailing their predicted chance of study success in the first year of their enrolled study program, an overview of their competences compared to peers, and links to remediation activities. Furthermore, the students are also surveyed in several areas about these feedback reports (e.g., feedback self-efficacy) (Fonteyne, 2017, 2022; Schelfhout, 2019). As of February 2019, a third application, ‘SIMON zegt het opnieuw’, offers additional support to first-year Ghent University students after their first exam period. Depending on their predicted chance of study success (from ‘SIMON zegt’) and study progress, students are encouraged to reflect on their study choice, referred to remediation initiatives or counseling advisors, or congratulated on their performance (Fonteyne, 2022). Both ‘Vraag het aan SIMON’ and ‘SIMON zegt’ consist of two primary components: the interest component (SIMON-I) and the competence component (SIMON-C).

First, SIMON-I measures vocational interests using the RIASEC model (Holland, 1997). Vocational interests encompass an individual's preferences towards specific types of work-related activities and environments (Rounds & Su, 2014; Stoll et al., 2017). Holland’s

RIASEC model (1997) helps categorizing and comprehending these vocational interests by defining six vocational interest dimensions: Realistic (R), Investigative (I), Artistic (A), Social (S), Enterprising (E), and Conventional (C). The establishment of the study programs interest profiles is based on RIASEC data from successful third-year bachelor's and first-year master's students. This approach allows for matching a (prospective) student's interest profile with a study program interest profile. The congruence between the interest profiles of an individual and an environment is known as person-environment (PE) interest fit (Hoff et al., 2020; Nye, Prasad, et al., 2018). To calculate PE interest fit in SIMON, Fonteyne (2017) used the method of checking the alignment between the top letters (e.g., top three) of the RIASEC profiles. Also, SIMON-I includes a subscale to assess specific interests in either academically or professionally oriented study programs (Fonteyne, 2017; Fonteyne, Wille, et al., 2017; Schelfhout, 2019). Upon finishing SIMON-I, (prospective) students receive their percentage scores on the six RIASEC dimensions and the academic versus professional scale, along with an overview of how well their interests match with various study programs (Fonteyne, 2017; Schelfhout, 2019).

Second, SIMON-C assesses competencies through a test battery that includes background surveys (including e.g., education type secondary education) and validated tests and questionnaires, measuring non-cognitive (e.g., self-efficacy, test anxiety) and cognitive characteristics (e.g., basic mathematical proficiency) known to predict academic achievement (Fonteyne, 2017; Fonteyne, Duyck, et al., 2017; Schelfhout, 2019). After completing the test battery, (prospective) students receive a predicted personal chance of study success. Specifically, prospective students get program-specific chances of study success for different study programs. First-year students are presented with the chance of study success in their chosen study program and a competence overview, compared to peers and supplemented with remediation initiatives in a personalized feedback report (Fonteyne, 2017, 2022). Importantly, SIMON-C focuses on achieving high accuracy in its predictions and identifying students who are very likely to lack the essential skills needed to successfully finish the first year of their studies (Fonteyne, 2017; Schelfhout, 2019). These prediction models are refined annually using recursive feature elimination and cross-validation, already drawing on data from tens of thousands of students (Fonteyne, 2017; Fonteyne, Duyck, et al., 2017; Schelfhout, 2019). Initially, the chances of study success focused solely on first-year academic achievement, but they now also project the likelihood of obtaining a bachelor's degree within three or four years (Fonteyne, 2022).

Overall, the ultimate goal of SIMON is to improve students' higher education journeys, especially for those at risk of failure. The present dissertation aims to optimize SIMON and thereby contribute to research and educational practice in multiple areas. Consequently, SIMON-C (Chapter 2 and Chapter 3), the corresponding feedback module for first-year students after participating in the 'SIMON zegt' application (Chapter 4) as well as SIMON-I (Chapter 5) are considered.

Competences

To provide students with insights into their personal and program-specific chances of study success in higher education through SIMON, previous research within the SIMON project examined the predictive validity of various background and (non-)cognitive student characteristics (i.e., SIMON-C) for academic achievement (Fonteyne, 2017; Fonteyne, Duyck, et al., 2017; Fonteyne et al., 2015). The combination of background factors (e.g., hours mathematics in secondary education), cognitive skills (e.g., basic mathematical proficiency), and non-cognitive characteristics (e.g., motivation and test anxiety) is found to be predictive for academic achievement and accounts for on average 23% of its variance. Furthermore, program-specific predictions identify 10% more at-risk students compared to general predictions across all study programs (Fonteyne, Duyck, et al., 2017). Additionally, regarding the specific cognitive skills, SIMON-C measures basic mathematical proficiency using a specially designed and validated basic mathematics test, while vocabulary and English reading comprehension are assessed with pre-existing validated tests (Fonteyne et al., 2015). However, Dutch academic language proficiency in terms of reading comprehension is not included in previous SIMON research due to the lack of a suitable and validated test at the time. Academic language proficiency aligns with Hulstijn's (2015) concept of Higher Language Cognition (HLC) compared with Basic Language Cognition (BLC). While BLC covers everyday language skills, HLC encompasses more advanced lexical, syntactical, and cognitive language abilities. Additionally, reading comprehension is a component of academic language proficiency and can be defined as understanding, using, evaluating, and reflecting (on) textual content (OECD, 2023). As articulated by Kintsch's (2013) Construction-Integration model, reading comprehension involves constructing meaning from this textual content and integrating it with existing long-term knowledge. This process requires the coordination of lower-order skills like decoding with higher-order skills such as vocabulary knowledge. In fact, academic language proficiency in the language of instruction is a known predictor of academic achievement (Elder, 2017; Heeren et al., 2021; Knoch & Elder, 2013).

The present dissertation further elucidates the refinement of SIMON-C's prediction models by first considering the occurrence of the COVID-19 pandemic. Indeed, this unexpected crisis emerged in the spring of 2019-2020 and acknowledging its impacts on academic achievement is relevant for the further implementation of SIMON. Second, we investigate Dutch academic language proficiency in terms of reading comprehension as an additional cognitive student characteristic in SIMON-C.

COVID-19 Pandemic

The COVID-19 pandemic has triggered substantial disruptions in global education, affecting 94% of students worldwide (United Nations, 2020). Educational institutions swiftly shifted to (partial) distance learning to ensure continuity in teaching and learning (Donnelly & Patrinos, 2021; Iterbeke & De Witte, 2021). Concerns have arisen regarding the pandemic's impact on academic achievement and potential learning losses (Azevedo et al., 2022; OECD, 2021b), particularly among students from low socio-economic backgrounds (Betthäuser et al., 2023; Moscoviz & Evans, 2022).

While (partial) distance learning shows potential to positively impact academic achievement under normal circumstances (for meta-analyses, see Bernard et al., 2014; Means et al., 2013; Vo et al., 2017), the pandemic necessitated an abrupt transition to (partial) distance learning (Adedoyin & Soykan, 2020; OECD, 2021b). Most research on pandemic-related learning losses focused on compulsory education but limited attention is given to higher education. A recent systematic review and meta-analysis of 42 studies in compulsory education across 15 countries indicates student learning losses with an overall effect size of Cohen's d = -0.14 (Betthäuser et al., 2023). In higher education, few studies comparing pre-pandemic and pandemic learning and using an objective outcome measure show inconsistent findings. Some of them report learning losses (Bird et al., 2022; De Paolo et al., 2022; Orlov et al., 2021), whereas others find no differences (El Said, 2021) or even improvements in academic achievement (Gonzalez et al., 2020; Iglesias-Pradas et al., 2021; Rodríguez-Planas, 2022). Crucially, these studies generally focused on early pandemic impacts.

Furthermore, student characteristics, including socio-economic status (SES), sex, cognitive ability, and non-cognitive factors (e.g., motivation, self-control, academic self-efficacy, and test anxiety), play a role in predicting academic achievement (Azevedo et al., 2022; Voyer & Voyer, 2014). However, possible interactions between the pandemic and, in particular, (non-)cognitive student characteristics on academic achievement remain understudied (Betthäuser et al., 2023; Iterbeke & De Witte, 2021).

Therefore, in Chapter 2, COVID-19 pandemic's effects on academic achievement in higher education are examined from a macro-level perspective, including four pre-pandemic years, a one-third pandemic year and a full pandemic year. Particular attention is also given to the potential moderating roles of background/(non-)cognitive student characteristics in the relationship between the pandemic and academic achievement.

Dutch Language Proficiency

The 21st-century society demands adequate language proficiency from individuals for full participation (Kennedy & Sundberg, 2020; OECD, 2019). In addition, academic language proficiency (hereafter: language proficiency) is crucial for higher education students (Knoch & Elder, 2013; Read, 2016). Specifically, reading comprehension is a component of language proficiency that poses challenges for students early in their higher education journey (De Wachter et al., 2013; Jansen et al., 2022; Van Houtven et al., 2010).

Language proficiency assessments serve admission purposes, mostly for non-native English speakers (Abunawas, 2014; Gagen, 2019; Ihlenfeldt & Rios, 2023; Wongtrirat, 2010), or post-entry screening by higher education institutions with low or no admission requirements to identify at-risk students for further support (Elder, 2017; Knoch & Elder, 2013; Read, 2016). Overall, these assessments show that language proficiency predicts academic achievement, explaining up to 10% of its variance (Elder, 2017; Heeren et al., 2021; Knoch & Elder, 2013). To our knowledge, however, a scarcity of studies reports on the predictive validity of post-entry language assessments (PELAs) for academic achievement (e.g., De Wachter et al., 2013; Heeren et al., 2021; van Dijk, 2015). Additionally, PELAs for languages other than English are currently highly limited.

Further, a recurring issue with internally developed PELAs is the absence of professional construct validation (Knoch & Elder, 2013), a trend noted across different language assessment studies as well (Min & Aryadoust, 2021). Moreover, the evaluation of language proficiency's predictive validity for academic achievement often overlooks study program-specific analyses. Nonetheless, this approach can elucidate how language proficiency impacts academic achievement within distinct fields of study (Elder, 2017; Hauspie et al., 2024; Read, 2016). When specifically considering a notable example for screening first-year university students in an open access higher education system (Heeren et al., 2021), we also observe that the test and its detailed validation evidence are not publicly accessible, affecting transparency, replicability, and reproducibility (Min & Aryadoust, 2021).

In Chapter 3, we validate the Ghent University Language Screening (GULS), developed as a Dutch PELA. GULS is a fully open access test (i.e., easy to administer, free and publicly accessible) and specifically assesses reading comprehension of first-year higher education students. Firstly, we investigate GULS's construct validity at both the model and item levels, along with its reliability, with data from two consecutive three-year periods. Secondly, we examine GULS's predictive validity for academic achievement, including Grade Point Average (GPA) and study success, using data from the same two three-year periods across 16 bachelor’s study programs, and for each study program separately across the six-year period.

Feedback

Through ‘SIMON zegt’, students receive feedback on their predicted chance of study success in the first year of their enrolled higher education study program, their competencies relative to peers, and available remediation/competence training initiatives. Feedback is information that can be derived from various sources, including teachers, students, and computer-based systems, and offers students insights into their current performance, goal alignment, future direction, and improvement suggestions (Lipnevich & Panadero, 2021). However, for feedback to be effective in improving learning outcomes like academic achievement, students need to engage with this feedback meaningfully. Research within the SIMON project reveals that negative feedback on attainability triggers both goal engagement (assimilation) and disengagement (accommodation), with the perceived accuracy of the feedback playing an essential role. Additionally, lower levels of motivation are associated with reduced goal engagement and increased disengagement, whereas academic self-efficacy does not influence these outcomes (Fonteyne et al., 2018). To further establish effective interventions within SIMON that encourage student feedback engagement, more research into its determinants and mechanisms is necessary, preferably through experimental and/or longitudinal studies.

Student Feedback Engagement

Feedback is recognized as a fundamental pillar of educational practice and policy (Hattie & Timperley, 2007; Panadero, 2023), with research highlighting its multifaceted nature and different impacts on diverse student outcomes (Kluger & DeNisi, 1996; Wisniewski et al., 2020). While much attention has been devoted to feedback *provision* (e.g., feedback type) (Wisniewski et al., 2020), understanding feedback *reception* is equally important (Boud & Molloy, 2013; Van der Kleij & Lipnevich, 2020). Indeed, the approach and the extent to which students engage with received feedback is essential to ensure that feedback effectively influences student outcomes (Ajjawi & Boud, 2017; Van der Kleij & Lipnevich, 2020). This acknowledgement aligns with the socio-constructivist feedback view that emphasizes the active role of learners in the feedback process (Dann, 2017; Winstone et al., 2019).

Despite the growing focus on feedback reception, there persists a necessity for studies on student feedback engagement and its underlying mechanisms, particularly concerning student characteristics as determinants (Panadero & Lipnevich, 2022; Winstone & Nash, 2023). Indeed, nuanced insights into student feedback engagement help find strategies to empower students in maximizing the feedback benefits for their learning journey. To achieve this goal, applying the Theory of Planned Behavior (TPB) (Ajzen, 1991, 2012) becomes relevant.

The TPB is renowned for its ability to predict and explain various behaviors, with an important role for behavioral intention influenced by factors such as perceived behavioral control (for meta-analyses, see e.g., Hirschey et al., 2020; Riebl et al., 2015). Widely accepted and used across various domains like business and health (Bosnjak et al., 2020), the model's implementation is gaining traction in educational contexts as well (e.g., Opoku et al., 2021). This framework enables the differentiation between intentional and behavioral feedback engagement, with feedback self-efficacy serving as the equivalent of perceived behavioral control. Intentional feedback engagement reflects an individual's willingness to engage with received feedback, signifying their readiness to invest time and effort. Conversely, behavioral feedback engagement represents active feedback engagement, wherein individuals take tangible actions based on the feedback (Ellis, 2010; Handley et al., 2011; Yu et al., 2019). Feedback self-efficacy, in turn, pertains to students' confidence in their ability to engage with received feedback effectively (Linderbaum & Levy, 2010; Winstone et al., 2019). This student characteristic influences student feedback engagement (Handley et al., 2011) and can be shaped by various information-related characteristics like feedback (Ajzen, 2020; Ajzen & Fishbein, 2005). Building upon the TPB, intentional feedback engagement also appears to play a mediating role in the relationship between feedback self-efficacy and behavioral feedback engagement (Ajzen, 1991, 2012), as well as feedback self-efficacy does in the relationship between feedback and intentional feedback engagement (Lipnevich & Panadero, 2021; Panadero, 2023).

Accordingly, in Chapter 4, a TPB-based model is considered to examine student feedback engagement among first-year university students who received feedback on their first-year predicted chance of study success (i.e., (very) low, (fairly) high), supplemented with recommendations for remediation/competence training activities. More specifically, the effects of intentional feedback engagement on behavioral feedback engagement, feedback self-efficacy on intentional and behavioral feedback engagement, and feedback on feedback self-efficacy are examined. Furthermore, attention is directed towards the mediations of intentional feedback engagement in the relationship between feedback self-efficacy and behavioral feedback engagement, and of feedback self-efficacy in the relationship between feedback and intentional feedback engagement.

Interests

To offer students an understanding of how their interest profile matches with the interest profiles of higher education study programs, research during the initial phase of the SIMON project developed the vocational interest questionnaire and assessed its validity and practical utility. Notably, students demonstrate a positive reaction to their interest profiles and the matched study programs (Fonteyne, Wille, et al., 2017). Another SIMON study shows that female students tend to have a better PE interest fit with their chosen STEM or non-STEM programs than male students, among other things (Schelfhout et al., 2021). Furthermore, previous SIMON investigations with a focus on vocational interests also proposed methodological refinements for the tool's optimization. For example, Schelfhout (2019) suggests an approach to identify students at risk of failure in their chosen programs using interests and a set of background, cognitive, and non-cognitive factors. Vocational interests are found to feature in 24% of the included predictive models for academic achievement, ranking as the third most crucial predictor following study antecedents and cognitive ability (Schelfhout, 2019). Regarding the calculation of the fit between a student's and a study program's interest profiles specifically, previous research introduced regressed interest fit for a single study environment, which demonstrates a positive relationship with academic achievement (r = .36) (Schelfhout et al., 2022). Studies of this more methodological nature are highly relevant for advancing empirical research and the practical implementation of PE interest fit in counseling settings such as SIMON.

Logistic Regressed Person-Environment Interest Fit

At the heart of PE fit theories lies the concept of vocational interests (Rounds & Su, 2014; Stoll et al., 2017), often categorized and understood through influential frameworks like Holland's RIASEC model (1997). By completing a RIASEC interest questionnaire, an individual's profile across these interest dimensions can be assessed, providing insights into their vocational preferences (Holland, 1997; Schelfhout et al., 2019). To determine an environment's interest profile, the incumbent method can be used. This method involves averaging the scores of individuals already enrolled in a specific environment (i.e., incumbents), such as a study program (Allen & Robbins, 2010; B. Schneider, 1987). Once the interest profiles of an individual and an environment are known, the calculation of PE interest fit is possible (Hoff et al., 2020; Nye, Prasad, et al., 2018).

In studying PE interest fit, early approaches relied on congruence indices to match dominant interest dimensions between individuals and environments (Edwards, 1993; Tinsley, 2000). The field progressed to embrace continuous PE interest fit methods like correlation fit and Euclidean distance fit (Tracey et al., 2012; Wille et al., 2014). Correlation fit involves calculating the correlation between an individual's RIASEC scores and those of the environment (Allen & Robbins, 2010; Tracey et al., 2012). Meanwhile, Euclidean distance reduces the person and environment interest profiles into two points in Euclidean space, with greater congruence indicated by points closer together (Wilkins & Tracey, 2014; Wille et al., 2014). More recently, regression-based approaches have gained traction.

Regressed PE interest fit involves regressing a criterion such as academic achievement on individuals' RIASEC scores to determine PE interest fit (Nye, Prasad, et al., 2018; Schelfhout et al., 2022). This approach assesses PE interest fit more precisely and better predicts academic achievement (Nye, Prasad, et al., 2018; Schelfhout et al., 2022) and work satisfaction (Nye, Butt, et al., 2018; Nye, Prasad, et al., 2018) than traditional methods. Furthermore, regressed PE interest fit offers a time-efficient advantage over these more conventional PE interest fit methods by eliminating the need for additional data collection to obtain environment interest profiles (Edwards, 1994; Xu & Li, 2020).

Notwithstanding these advancements, certain gaps persist in literature. While much attention is paid to the prediction of academic achievement by regressed PE interest fit (Nye, Butt, et al., 2018; Nye et al., 2012; Schelfhout et al., 2022), comparatively little emphasis is placed on the equally important aspect of study choice. Indeed, vocational interests are considered one of the best predictors of study choice (Rounds & Su, 2014; Stoll et al., 2017). Moreover, PE interest fit methods overlook the fact that individuals not only seek environments that align with their vocational interests but also actively avoid those that do not (De Cooman & Vleugels, 2022; Holland, 1997; B. Schneider, 1987). For example, when choosing a study program like Psychology, students compare themselves with both the interest patterns that *align* with Psychology as well as with the interest patterns that *divert* from Psychology (Feldman et al., 2001; Weidman, 2005). Incorporating this interest pattern differentiation into PE interest fit measures acknowledges students' exploratory approach to study choice and facilitates improved support.

In Chapter 5, we introduce Logistic Regressed PE Interest Fit (LRIF). LRIF differentiates between interest patterns that align with and those that divert from a specific environment like a study program, without the need for additional data collection to predetermine environment interest profiles. First, LRIF's predictive validity for study choice is evaluated, considering 31 study programs. Second, LRIF is compared with more traditional continuous PE interest fit measures (i.e., correlation and Euclidean distance fit) in predicting study choice.

References

Abunawas, M. (2014). *A meta-analytic investigation of the predictive validity of the Test of English as a Foreign Language (TOEFL) scores on GPA*. https://oaktrust.library.tamu.edu/bitstream/handle/1969.1/154156/ABUNAWAS-DISSERTATION-2014.pdf?sequence=1

Adedoyin, O. B., & Soykan, E. (2020). Covid-19 pandemic and online learning: the challenges and opportunities. *Interactive Learning Environments*, 1–13. https://doi.org/10.1080/10494820.2020.1813180

Ajjawi, R., & Boud, D. (2017). Researching feedback dialogue: an interactional analysis approach. *Assessment & Evaluation in Higher Education*, *42*(2), 252–265. https://doi.org/10.1080/02602938.2015.1102863

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Ajzen, I. (2012). The Theory of Planned Behavior. In *Handbook of Theories of Social Psychology: Volume 1* (pp. 438–459). SAGE Publications Ltd. https://doi.org/10.4135/9781446249215.n22

Ajzen, I. (2020). The theory of planned behavior: Frequently asked questions. *Human Behavior and Emerging Technologies*, *2*(4), 314–324. https://doi.org/10.1002/hbe2.195

Ajzen, I., & Fishbein, M. (2005). The Influence of Attitudes on Behavior. In D. Albarracin, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173–221). Lawrence Erlbaum Associates.

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. https://doi.org/10.1037/a0017267

Azevedo, P., Gutierrez, M., de Hoyos, R., & Saavedra, J. (2022). The Unequal Impacts of COVID-19 on Student Learning. In F. M. Reimers (Ed.), *Primary and Secondary Education During Covid-19. Disruptions to Educational Opportunity During a Pandemic* (pp. 421–459). https://doi.org/10.1007/978-3-030-81500-4_16

Bernard, R. M., Borokhovski, E., Schmid, R. F., Tamim, R. M., & Abrami, P. C. (2014). A meta-analysis of blended learning and technology use in higher education: from the general to the applied. *Journal of Computing in Higher Education*, *26*(1), 87–122. https://doi.org/10.1007/s12528-013-9077-3

Betthäuser, B. A., Bach-Mortensen, A. M., & Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. *Nature Human Behaviour*, *7*(3), 375–385. https://doi.org/10.1038/s41562-022-01506-4

Bird, K. A., Castleman, B. L., & Lohner, G. (2022). Negative Impacts From the Shift to Online Learning During the COVID-19 Crisis: Evidence From a Statewide Community College System. *AERA Open*, *8*, 233285842210812. https://doi.org/10.1177/23328584221081220

Bosnjak, M., Ajzen, I., & Schmidt, P. (2020). The theory of planned behavior: Selected recent advances and applications. *Europe's Journal of Psychology*, *16*(3), 352–356. https://doi.org/10.5964/ejop.v16i3.3107

Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assessment & Evaluation in Higher Education*, *38*(6), 698–712. https://doi.org/10.1080/02602938.2012.691462

Dann, R. (2017). *Developing feedback for pupil learning: Teaching, learning and assessment in schools.* Routledge.

De Cooman, R., & Vleugels, W. (2022). Person–Environment Fit: Theoretical Perspectives, Conceptualizations, and Outcomes. In *Oxford Research Encyclopedia of Business and Management*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190224851.013.377

De Paolo, M., Gioia, F., & Scoppa, V. (2022). *Online Teaching, Procrastination and Students' Achievement: Evidence from COVID-19 Induced Remote Learning*.

De Wachter, L., Heeren, J., Marx, S., & Huyghe, S. (2013). "Taal: een noodzakelijke, maar niet de enige voorwaarde tot studiesucces: De correlatie tussen de resultaten van een taalvaardigheidstoets en de slaagcijfers bij eerstejaarsstudenten aan de KU Leuven." [Language Proficiency: A Necessary, but Not the Only, Condition for Study Success: A Correlation between the Results of a Language Proficiency Test and Academic Achievement of First-Year Students.]. *Levende Talen Tijdschrift*, *14*(4), 28–36. https://lt-tijdschriften.nl/ojs/index.php/ltt/article/view/549

Donnelly, R., & Patrinos, H. A. (2021). Learning loss during Covid-19: An early systematic review. *PROSPECTS*. https://doi.org/10.1007/s11125-021-09582-6

Edwards, J. R. (1993). Problems with the Use of Profile Similarity Indices in the Study of Congruence in Organizational Research. *Personnel Psychology*, *46*(3), 641–665. https://doi.org/10.1111/j.1744-6570.1993.tb00889.x

Edwards, J. R. (1994). The Study of Congruence in Organizational Behavior Research: Critique and a Proposed Alternative. *Organizational Behavior and Human Decision Processes*, *58*(1), 51–100. https://doi.org/10.1006/obhd.1994.1029

El Said, G. R. (2021). How Did the COVID-19 Pandemic Affect Higher Education Learning Experience? An Empirical Investigation of Learners' Academic Performance at a

University in a Developing Country. *Advances in Human-Computer Interaction*, *2021*, 1–10. https://doi.org/10.1155/2021/6649524

Elder, C. (2017). Language Assessment in Higher Education. In *Language Testing and Assessment* (pp. 271–286). Springer International Publishing. https://doi.org/10.1007/978-3-319-02261-1_35

Ellis, R. (2010). A Framework for Investigating Oral and Written Corrective Feedback. *Studies in Second Language Acquisition*, *32*(2), 335–349. https://doi.org/10.1017/S0272263109990544

European Commission. (2021, January 26). *European Credit Transfer and Accumulation System (ECTS).* https://ec.europa.eu/education/resources-and-tools/european-credit-transfer-and-accumulation-system-ects_en

Eurydice. (2023, November 27). *Belgium - Flemish Community*. https://eurydice.eacea.ec.europa.eu/national-education-systems/belgium-flemish-community/bachelor

Feldman, K. A., Ethington, C. A., & Smart, J. C. (2001). A Further Investigation of Major Field and Person-Environment Fit. *The Journal of Higher Education*, *72*(6), 670–698. https://doi.org/10.1080/00221546.2001.11777121

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.

Fonteyne, L. (2022). SIMON biedt handvatten voor studiekeuze en -succes. *TH&MA (DEN HAAG)*, *2022*(4), 14–18. http://hdl.handle.net/1854/LU-01GXAP2CGKG5SKH968VKDG4F3A

Fonteyne, L., De Fruyt, F., Dewulf, N., Duyck, W., Erauw, K., Goeminne, K., Lammertyn, J., Marchant, T., Moerkerke, B., Oosterlinck, T., & Rosseel, Y. (2015). Basic mathematics test predicts statistics achievement and overall first year academic success. *European Journal of Psychology of Education*, *30*(1), 95–118. https://doi.org/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. https://doi.org/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. https://doi.org/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 for Educational and Vocational Guidance*, *17*(2), 233–262. https://doi.org/10.1007/s10775-016-9327-9

Gagen, T. (2019). *The predictive validity of IELTS scores: A meta-analysis*. https://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=8762&context=etd

Gonzalez, T., de la Rubia, M. A., Hincz, K. P., Comas-Lopez, M., Subirats, L., Fort, S., & Sacha, G. M. (2020). Influence of COVID-19 confinement on students' performance in higher education. *PLOS ONE*, *15*(10), e0239490. https://doi.org/10.1371/journal.pone.0239490

Handley, K., Price, M., & Millar, J. (2011). Beyond 'doing time': investigating the concept of student engagement with feedback. *Oxford Review of Education*, *37*(4), 543–560. https://doi.org/10.1080/03054985.2011.604951

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hauspie, C., Schelfhout, S., Dirix, N., Fonteyne, L., Janse, M., Szmalec, A., Vereeck, A., & Duyck, W. (2024). Does Studying Latin in Secondary Education Predict Study Achievement in Academic Higher Education? *Language Learning*. https://doi.org/10.1111/lang.12639

Heeren, J., Speelman, D., & De Wachter, L. (2021). A practical academic reading and vocabulary screening test as a predictor of achievement in first-year university students: implications for test purpose and use. *International Journal of Bilingual Education and Bilingualism*, *24*(10), 1458–1473. https://doi.org/10.1080/13670050.2019.1709411

Hirschey, R., Bryant, A. L., Macek, C., Battaglini, C., Santacroce, S., Courneya, K. S., Walker, J. S., Avishai, A., & Sheeran, P. (2020). Predicting physical activity among cancer survivors: Meta-analytic path modeling of longitudinal studies. *Health Psychology*, *39*(4), 269–280. https://doi.org/10.1037/hea0000845

Hoff, K. A., Song, Q. C., Wee, C. J. M., Phan, W. M. J., & Rounds, J. (2020). Interest fit and job satisfaction: A systematic review and meta-analysis. *Journal of Vocational Behavior*, *123*, 103503. https://doi.org/10.1016/j.jvb.2020.103503

Holland, J. L. (1997). *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments (3rd ed.)*. Psychology Assessment Resources.

Hulstijn, J. H. (2015). *Language Proficiency in Native and Non-native Speakers*. John Benjamins B.V.

Iglesias-Pradas, S., Hernández-García, Á., Chaparro-Peláez, J., & Prieto, J. L. (2021). Emergency remote teaching and students' academic performance in higher education during the COVID-19 pandemic: A case study. *Computers in Human Behavior*, *119*, 106713. https://doi.org/10.1016/j.chb.2021.106713

Ihlenfeldt, S. D., & Rios, J. A. (2023). A meta-analysis on the predictive validity of English language proficiency assessments for college admissions. *Language Testing*, *40*(2), 276–299. https://doi.org/10.1177/02655322221112364

Iterbeke, K., & De Witte, K. (2021). Helpful or Harmful? The Role of Personality Traits in Student Experiences of the COVID-19 Crisis and School Closure. *Personality and Social Psychology Bulletin*, 014616722110505. https://doi.org/10.1177/01461672211050515

Jansen, C., De Wachter, L., Van Dun, P., & Frik, T. (2022). *Taalcompetentie in het Nederlands van Nederlandstalige studenten in het hoger onderwijs in Nederland en Vlaanderen*. https://taalunie.org/publicaties/213/taalcompetentie-in-het-nederlands-van-nederlandstalige-studenten-in-het-hoger-onderwijs-in-nederland-en-vlaanderen

Kelchtermans, S., & Verboven, F. (2010). Participation and study decisions in a public system of higher education. *Journal of Applied Econometrics*, *25*(3), 355–391. https://doi.org/10.1002/jae.1087

Kennedy, T. J., & Sundberg, C. W. (2020). *21st Century Skills* (pp. 479–496). https://doi.org/10.1007/978-3-030-43620-9_32

Kintsch, W. (2013). Revisiting the Construction–Integration Model of Text Comprehension and Its Implications for Instruction. In *Theoretical Models and Processes of Reading* (pp. 807–839). International Reading Association. https://doi.org/10.1598/0710.32

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284. https://doi.org/10.1037/0033-2909.119.2.254

Knoch, U., & Elder, C. (2013). A framework for validating post-entry language assessments (PELAs). *Papers in Language Testing and Assessment*, *2*(2), 48–66. https://arts.unimelb.edu.au/__data/assets/pdf_file/0007/1771306/4_Knoch_Elder_0.pdf

Linderbaum, B. A., & Levy, P. E. (2010). The Development and Validation of the Feedback Orientation Scale (FOS). *Journal of Management*, *36*(6), 1372–1405. https://doi.org/10.1177/0149206310373145

Lipnevich, A. A., & Panadero, E. (2021). A Review of Feedback Models and Theories: Descriptions, Definitions, and Conclusions. *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.720195

Min, S., & Aryadoust, V. (2021). A systematic review of item response theory in language assessment: Implications for the dimensionality of language ability. *Studies in Educational Evaluation*, *68*, 100963. https://doi.org/10.1016/j.stueduc.2020.100963

Moscoviz, L., & Evans, D. K. (2022). *Learning Loss and Student Dropouts during the COVID-19 Pandemic: A Review of the Evidence Two Years after Schools Shut Down*. https://www.cgdev.org/sites/default/files/learning-loss-and-student-dropouts-during-covid-19-pandemic-review-evidence-two-years.pdf

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. https://doi.org/10.1016/j.jvb.2018.08.003

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. https://doi.org/10.1016/j.jvb.2017.10.012

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational Interests and Performance. *Perspectives on Psychological Science*, *7*(4), 384–403. https://doi.org/10.1177/1745691612449021

OECD. (2019). *An OECD Learning Framework 2030* (pp. 23–35). https://doi.org/10.1007/978-3-030-26068-2_3

OECD. (2021a). *Resourcing Higher Education in the Flemish Community of Belgium, Higher Education*. https://doi.org/https://doi.org/10.1787/3f0248ad-en

OECD. (2021b). *The State of Global Education - 18 Months into the Pandemic*. https://doi.org/10.1787/1a23bb23-en

OECD. (2022). *Education at a Glance 2022*. OECD. https://doi.org/10.1787/3197152b-en

OECD. (2023). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. https://doi.org/https://doi.org/10.1787/53f23881-en

Opoku, M. P., Cuskelly, M., Pedersen, S. J., & Rayner, C. S. (2021). Applying the theory of planned behaviour in assessments of teachers' intentions towards practicing inclusive education: a scoping review. *European Journal of Special Needs Education*, *36*(4), 577–592. https://doi.org/10.1080/08856257.2020.1779979

Oppedisano, V. (2009). *Open University Admission Policies and Drop Out Rates in Europe*. https://www.researchgate.net/publication/4646637_Open_University_Admission_Policies_and_Drop_Out_Rates_in_Europe

Orlov, G., McKee, D., Berry, J., Boyle, A., DiCiccio, T., Ransom, T., Rees-Jones, A., & Stoye, J. (2021). Learning during the COVID-19 pandemic: It is not who you teach, but how you teach. *Economics Letters*, *202*, 109812. https://doi.org/10.1016/j.econlet.2021.109812

Panadero, E., & Lipnevich, A. A. (2022). A review of feedback models and typologies: Towards an integrative model of feedback elements. *Educational Research Review*, *35*, 100416. https://doi.org/10.1016/j.edurev.2021.100416

Pokhrel, S., & Chhetri, R. (2021). A Literature Review on Impact of COVID-19 Pandemic on Teaching and Learning. *Higher Education for the Future*, *8*(1), 133–141. https://doi.org/10.1177/2347631120983481

Read, J. (2016). *Some Key Issues in Post-Admission Language Assessment* (pp. 3–20). https://doi.org/10.1007/978-3-319-39192-2_1

Riebl, S. K., Estabrooks, P. A., Dunsmore, J. C., Savla, J., Frisard, M. I., Dietrich, A. M., Peng, Y., Zhang, X., & Davy, B. M. (2015). A systematic literature review and meta-analysis: The Theory of Planned Behavior's application to understand and predict nutrition-related behaviors in youth. *Eating Behaviors*, *18*, 160–178. https://doi.org/10.1016/j.eatbeh.2015.05.016

Rodríguez-Planas, N. (2022). COVID-19, college academic performance, and the flexible grading policy: A longitudinal analysis. *Journal of Public Economics*, *207*, 104606. https://doi.org/10.1016/j.jpubeco.2022.104606

Rounds, J., & Su, R. (2014). The Nature and Power of Interests. *Current Directions in Psychological Science*, *23*(2), 98–103. https://doi.org/10.1177/0963721414522812

Schelfhout, S. (2019). *The Influence of Study Interests and (Non-)Cognitive Predictors on Study Choice and Study Achievement in Flemish Higher Education* [Dissertation]. Ghent University.

Schelfhout, S., Bassleer, M., Wille, B., Van Cauwenberghe, S., Dutry, M., Fonteyne, L., Dirix, N., Derous, E., De Fruyt, F., & Duyck, W. (2022). Regressed person-environment interest fit: Validating polynomial regression for a specific environment. *Journal of Vocational Behavior*, *136*, 103748. https://doi.org/10.1016/j.jvb.2022.103748

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

Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., Derous, E., De Fruyt, F., & Duyck, W. (2021). How interest fit relates to STEM study choice: Female students fit their choices better. *Journal of Vocational Behavior*, *129*, 103614. https://doi.org/10.1016/j.jvb.2021.103614

Schleicher, A. (2020). *The impact of COVID-19 on education - Insights from Education at a Glance 2020*. https://www.oecd.org/education/the-impact-of-covid-19-on-education-insights-education-at-a-glance-2020.pdf

Schleicher, A. (2021). *The state of higher education. One year into the COVID-19 pandemic*. https://doi.org/10.1787/83c41957-en

Schneider, B. (1987). THE PEOPLE MAKE THE PLACE. *Personnel Psychology*, *40*(3), 437–453. https://doi.org/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. https://doi.org/10.1037/bul0000098

Statistiek Vlaanderen. (2024a, March 26). *Studieduur in het hoger onderwijs*. https://www.vlaanderen.be/statistiek-vlaanderen/onderwijs-en-vorming/studieduur-in-het-hoger-onderwijs

Statistiek Vlaanderen. (2024b, March 26). *Studierendement in het hoger onderwijs*. https://www.vlaanderen.be/statistiek-vlaanderen/onderwijs-en-vorming/studierendement-in-het-hoger-onderwijs

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. https://doi.org/10.1037/pspp0000117

Study in Flanders. (2024). *Tuition Fees*. https://www.studyinflanders.be/practical-information/tuition-fees

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. https://doi.org/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. https://doi.org/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. https://doi.org/10.1016/j.jvb.2011.03.005

United Nations. (2020). *Education during COVID-19 and beyond*. https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2020/08/sg_policy_brief_covid-19_and_education_august_2020.pdf

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

Van der Kleij, F. M., & Lipnevich, A. A. (2020). Student perceptions of assessment feedback: a critical scoping review and call for research. *Educational Assessment, Evaluation and Accountability*, *33*(2), 345–373. https://doi.org/10.1007/s11092-020-09331-x

van Dijk, T. (2015). Tried and tested. *Tijdschrift Voor Taalbeheersing*, *37*(2), 159–186. https://doi.org/10.5117/TVT2015.2.VAND

Vlaamse overheid. (2024). *Toelatingsexamens Vlaanderen*. https://toelatingsexamensvlaanderen.be/

Vo, H. M., Zhu, C., & Diep, N. A. (2017). The effect of blended learning on student performance at course-level in higher education: A meta-analysis. *Studies in Educational Evaluation*, *53*, 17–28. https://doi.org/10.1016/j.stueduc.2017.01.002

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174–1204. https://doi.org/10.1037/a0036620

Weidman, J. C. (2005). Academic Disciplines: Holland's Theory and the Study of College Students and Faculty. *The Journal of Higher Education*, *76*(2), 232–234. https://doi.org/10.1080/00221546.2005.11778912

Wilkins, K. G., & Tracey, T. J. G. (2014). Person Environment Fit and Vocational Outcomes. In *Psycho-social Career Meta-capacities* (pp. 123–138). Springer International Publishing. https://doi.org/10.1007/978-3-319-00645-1_7

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. https://doi.org/10.1016/j.jvb.2013.12.001

Winstone, N. E., Hepper, E. G., & Nash, R. A. (2019). Individual differences in self-reported use of assessment feedback: the mediating role of feedback beliefs. *Educational Psychology*, *41*(7), 844–862. https://doi.org/10.1080/01443410.2019.1693510

Winstone, N. E., & Nash, R. A. (2023). Toward a cohesive psychological science of effective feedback. *Educational Psychologist*, *58*(3), 111–129. https://doi.org/10.1080/00461520.2023.2224444

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.03087

Wongtrirat, R. (2010). *English language proficiency and academic achievement of international students: A meta-analysis*. https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1183&context=efl_etds

Xu, H., & Li, H. (2020). Operationalize Interest Congruence: A Comparative Examination of Four Approaches. *Journal of Career Assessment*, *28*(4), 571–588. https://doi.org/10.1177/1069072720909825

Yu, S., Zhang, Y., Zheng, Y., Yuan, K., & Zhang, L. (2019). Understanding student engagement with peer feedback on master's theses: a Macau study. *Assessment & Evaluation in Higher Education*, *44*(1), 50–65. https://doi.org/10.1080/02602938.2018.1467879

2

Effects of the COVID-19 Pandemic on Academic Achievement in Higher Education

Submitted to Educational Assessment, Evaluation and Accountability, 16th of January 2024.

Accepted for Revision, 7th of April 2024.

Mona Bassleer, Stijn Schelfhout, Lot Fonteyne, Maya Caen, Ilse De Bourdeaudhuij, Nicolas Dirix & Wouter Duyck.

Abstract

The emergence of the COVID-19 pandemic resulted in temporary closures of educational institutions and has shifted the educational process towards the use of (digital) distance education. Despite the efforts, severe learning losses and larger educational inequality are found in compulsory education. To complement this knowledge, the present prospective study focuses on higher education by analyzing academic achievement data of Flemish first-year university students (N = 24,404) spanning six years. COVID-19 learning losses are assessed in a natural setting, considering various background and (non-)cognitive student characteristics. Results for the full pandemic year 2020-2021 indicate that although the overall negative impact of the pandemic on academic achievement is rather small, the variance in academic achievement does increase. Low socio-economic status students show larger learning losses and the socio-economic achievement gap widens. Our findings imply that COVID-19 learning losses in higher education are less pervasive than in compulsory education, though inequality issues also arise.

Introduction

The COVID-19 pandemic has affected society globally across different areas. Educational institutions have faced the largest disruption in recent history, with closures that impacted 94% of the global studying population (United Nations, 2020). To guarantee teaching and learning continuity, a rapid and unprecedented transition towards (partial) distance learning was necessary (Donnelly & Patrinos, 2021; Iterbeke & De Witte, 2021). Concerns about this disruption's effect on students' academic achievement and related learning losses are voiced in literature (Azevedo et al., 2022; OECD, 2021). It is also suggested that learning losses are more severe among low socio-economic status (SES) students (Betthäuser et al., 2023; Moscoviz & Evans, 2022), which makes the pandemic a threat to global educational equality. However, existing research mainly focused on younger children and adolescents in compulsory education.

Currently, specific empirical research assessing pandemic learning losses and inequality in higher education is still scarce. Also, little is known about background/(non-)cognitive student characteristics other than SES and sex that may have influenced the relationship between the pandemic and academic achievement (Betthäuser et al., 2023; Iterbeke & De Witte, 2021). Studying these interactions is highly relevant as literature already indicates that individual differences in terms of cognitive ability and personality traits can have specific effects on learning outcomes (Azevedo et al., 2022; Voyer & Voyer, 2014).

The present study therefore investigates, from a macro-level perspective, how the COVID-19 pandemic affects academic achievement in higher education. Specific consideration is given to the suspected learning losses of low SES students and other determinants of academic achievement. By using a large dataset of six first-year university student cohorts between 2015 and 2021 (*N* = 24,404), derived from a running longitudinal project, the present study clarifies the scope and determinants of pandemic learning loss in higher education, informing educators, researchers and policymakers about targets of learning remediation (OECD, 2021; Pokhrel & Chhetri, 2021).

Impact of the Pandemic on (Higher) Education

The pandemic necessitated institutions to make a swift transition from traditional face-to-face learning to (partial) distance learning (i.e., online and blended learning) to ensure the continuity of high-quality education (Donnelly & Patrinos, 2021; Moscoviz & Evans, 2022). Blended learning can be defined as a combination of face-to-face and online teaching and learning (Means et al., 2013). The use of online or blended learning does not necessarily imply a problem. Meta-analytic evidence, including 45 (quasi-)experimental studies across educational stages, shows that academic achievement appears to be equivalent in the studies that compared purely online learning with face-to-face learning (Means et al., 2013), which is also confirmed in a recent meta-analysis focused on undergraduate medical students (Pei & Wu, 2019). However, some researchers also find lower academic achievement for online learning versus face-to-face learning in higher education (e.g., Bettinger et al., 2017). Meta-analyses involving studies that compared blended with face-to-face learning, of which two specifically focused on higher education (Bernard et al., 2014; Vo et al., 2017), indicate that academic achievement increases through blended learning (range mean Hedges' g [0.33, 0.39]) (e.g., Means et al., 2013). The extra learning time, additional instructional resources and interaction encouraging course elements that characterize good blended learning are put forward as possible determinants of learning gains (Means et al., 2013; Vo et al., 2017).

In general, a (partial) distance learning environment seems to be associated with positive learning outcomes. However, (partial) distance learning imposed by the pandemic has not been a typical case of planned and prepared (partial) distance learning, and often merely crisis management (Adedoyin & Soykan, 2020; OECD, 2021). The above meta-analyses only contain good practice implementations of (partial) distance learning on a small scale (e.g., in the context of single courses), whereas the pandemic now forced entire institutions and countries to shift to (partial) distance learning, without much preparation or a conceptual and didactic framework. And, of course, the context of the pandemic also implied many other (e.g., social) restrictions and health issues for students. The present study's goal is to assess the effects of the pandemic on academic achievement from a macro-level perspective (Betthäuser et al., 2023; OECD, 2021).

The question of learning losses following this crisis-response transition towards (partial) distance learning is relevant for all educational stages. However, the majority of research that specifically examined the influence of the COVID-19 pandemic on academic achievement is mainly focused on compulsory education (e.g., Engzell et al., 2021; Lichand et al., 2022). In minors, and, opposite to the positive effects of good practice (partial) distance learning (across educational stages) reported above, findings are disturbingly negative. For example, a recent systematic review and meta-analysis of 42 studies across 15 countries provide evidence of student learning losses with an overall Cohen's d = -0.14 (Betthäuser et al., 2023). In addition, on average across OECD countries, PISA[1] 2022 scores drop by about ten (in reading) to nearly fifteen (in mathematics) score points compared with PISA 2018 (i.e., d = 0.10 to d = 0.15 in the PISA distribution), which corresponds to a loss of one-half to three-fourths of a year of learning due to the pandemic (OECD, 2023).

To our knowledge, currently a few studies report data regarding the pandemic's impact on academic achievement in higher education by comparing pre-pandemic and pandemic learning and using an objective outcome measure. The results of the higher education pandemic studies are mixed, showing no academic achievement differences (El

[1] Programme for International Student Assessment. PISA tests 15-year-old students in three core domains (Mathematics, Reading and Science). The first assessment took place in 2000 and is managed every three years. Per cycle, one domain is considered the major one. A difference of ten PISA points is equivalent to an effect size d = 0.10 (Azevedo et al., 2020).

Said, 2021), learning losses (Bird et al., 2022; De Paolo et al., 2022; Orlov et al., 2021) and learning gains (Gonzalez et al., 2020; Iglesias-Pradas et al., 2021; Rodríguez-Planas, 2022). Importantly, all these studies contrasted academic achievement in academic year 2019-2020 with (a) previous academic year(s). The researchers thus only included a few months of (partial) distance learning, with lockdowns starting in March 2020. Because learning losses are likely to accumulate, the present study also focuses on academic year 2020-2021, which started on the last Monday in September and entirely took place in full pandemic. Also, we analyze a large sample, across study domains and a six-year period, in order to ensure generalizability and to control for normal fluctuations in academic achievement.

In sum, pre-pandemic meta-analytic evidence indicates that (partial) distance learning as such may benefit academic achievement (e.g., Means et al., 2013). The pandemic of course has much broader co-occurring negative effects, and most published COVID-19 studies report learning losses in compulsory education (Betthäuser et al., 2023; Moscoviz & Evans, 2022). To the best of our knowledge, no study compared academic achievement in (a) pre-pandemic year(s) and a full pandemic year in higher education. We expect to observe learning losses in the present study as well. Indeed, the pandemic is associated with a crisis-response migration to (partial) distance learning, concurrently with other pandemic restrictions and health concerns (Adedoyin & Soykan, 2020). A forced implementation of (partial) distance learning on a larger scale was thus unavoidable (OECD, 2021).

Role of Student Characteristics

Interactions of SES and other student characteristics with the pandemic's impact on academic achievement could (partially) help understand which types of students are more influenced by the pandemic (Kintu et al., 2017; Rodríguez-Hernández et al., 2020). Due to the unforeseeable nature of the pandemic outbreak, the present study includes the student characteristics that are already used in a longitudinal project linking study orientation with academic achievement. A unique opportunity presented itself to collect numerous student-level variables during the pandemic in an exceptionally large population.

Background Characteristics

Pre-pandemic meta-analyses on SES reveal a moderate to strong relation between SES and academic achievement in primary (Yu & Yu, 2021) and secondary education (Çiftçi & Cin, 2017), in favor of high SES students. The same applies for higher education, although the association is weaker here (Rodríguez-Hernández et al., 2020). In the (partial) distance learning context, high SES students also seem to benefit more from the mainly positive influence of (partial) distance learning on academic achievement (López-Pérez et al., 2011). Indeed, an economically and/or socially disadvantaged background can hinder the accessibility and affordability to e-learning infrastructures and the desired (parental) supportive environment for (partial) distance learning (OECD, 2021; Pokhrel & Chhetri, 2021). Meanwhile, systematic reviews show that, both in compulsory and higher education, the pandemic effect seems to interact with SES. Low SES students appear to have larger learning losses than high SES ones (Betthäuser et al., 2023; Donnelly & Patrinos, 2021; Moscoviz & Evans, 2022). PISA 2022 results also show that the SES (mathematics) academic achievement gap widens with seven score points (i.e., $d = 0.07$) compared with PISA 2018, averaged across OECD countries (OECD, 2023). Several underlying reasons can cause this gap. First, families with a high SES background are more likely to foresee (psychological) support, which is understood as especially relevant in a crisis situation (Hammerstein et al., 2021). Second, low SES students often experience more difficulty in obtaining access to technology needed for compensating the absent on campus student-teacher interactions (Azevedo et al., 2022; OECD, 2021). Third, lower SES seems to be associated with a higher risk of COVID-19 infection and mental distress (Anderson et al., 2020; Betthäuser et al., 2023), which could lead to lower study involvement. Therefore, in the present study, we expect to find lower academic achievement among low SES students compared with high SES ones. An additional negative impact of the pandemic, resulting in a wider socio-economic academic achievement gap, is hypothesized.

For sex, literature reports that females outperform males in academic achievement across different educational stages (Voyer & Voyer, 2014). Related to (partial) distance learning, a meta-analysis finds no sex differences in online learning outcomes (Yu & Yu, 2021), even though males appear to hold a more favorable attitude towards technology use than females (Cai et al., 2017). Across OECD countries and compared with 2018, PISA 2022 findings do reveal a widened sex (mathematics) academic achievement gap with four score points (i.e., $d = 0.04$) on average, in favor of males (OECD, 2023). However, the limited number of COVID-19 studies in higher education that included sex as a possible moderator shows no interaction between sex and the pandemic towards academic achievement (El Said, 2021; Orlov et al., 2021). Other research indicates that female versus male university students experience greater negative impacts from the COVID-19 pandemic in academics, struggling more with the shift to online learning and its effects on schoolwork. This trend also applies to perceived social isolation, stress, and mental health. Furthermore, frequent social media use as a coping mechanism during the pandemic intensifies the perceived negative effects on academics and stress for females, while affecting the perceived social relationships and mental health of females and males similarly (Prowse et al., 2021). In the present study, higher academic achievement for female students is hypothesized. Based on previous COVID-19 higher education studies (El Said, 2021; Orlov et al., 2021), we expect sex not to moderate the relationship between the pandemic and academic achievement.

(Non-)Cognitive Characteristics

Meta-analytic evidence emphasizes that cognitive ability is arguably one of the strongest predictive factors of academic achievement ($\rho = .54$) that increases throughout educational stages (Roth et al., 2015). In the (partial) distance learning context, most researchers only control for prior academic achievement, used as a proxy for cognitive ability, when explaining the variance in academic achievement (e.g., Vo et al., 2020). Some higher education studies also investigated the potential moderating role of prior academic achievement in the relationship between online (Bettinger et al., 2017) or blended learning (Asarta & Schmidt, 2017) and academic achievement. Their results indicate higher academic achievement in face-to-face versus blended learning for students with lower levels of prior academic achievement (Asarta & Schmidt, 2017). Bettinger and colleagues (2017) also discover larger negative effects of online learning on academic achievement among students with lower prior academic achievement. Some COVID-19 research, only in compulsory education, examined how academic achievement of low-, (average-) and high-achieving students, using population percentiles (Schult et al., 2021) or relative error rates (Spitzer & Musslick, 2021), differ between spring 2020 and at least one previous year. However, these studies did not include measures of (prior) academic achievement derived from (a) pretest(s) or academic achievement in one or more previous courses/educational stages. The present study contributes to fill this void by including prior academic achievement measures (i.e., hours of mathematics in secondary education and the secondary educational track), and language proficiency as proxies for cognitive ability. These factors are known as determinants of first-year academic achievement in higher education (Ashford et al., 2016; Heeren et al., 2021).

Besides cognitive characteristics, academic achievement is also influenced by non-cognitive socio-emotional skills and traits (Pierre et al., 2014), that are also assessed in our student sample. First, self-control can be described as the regulation of attentional, emotional and behavioral impulses to accomplish long-term goals (Duckworth et al., 2019). This characteristic has a positive impact on academic achievement across different ages, as reported in a systematic review (Duckworth et al., 2019). Self-control even seems to explain academic achievement above and beyond predictors such as cognitive ability (Stadler et al., 2016). In a (partial) distance learning environment, self-control also positively influences students' higher education academic achievement. This characteristic encompasses, among others, the ability to avoid distraction from interruption and using time effectively. But, when adding self-regulated learning and online engagement, these factors mediate the relation between self-control and academic achievement and the direct effect of self-control disappears (Zhu et al., 2016).

Second, motivation as a process of setting and striving for goals (Yu, 2021) can be distinguished in controlled (driven by extern factors) and autonomous (driven by internal factors) motivation (Deci & Ryan, 2008). Recent meta-analytic evidence, including both compulsory and higher education studies, indicates that improved academic achievement is mainly found in students with higher autonomous motivation (Howard et al., 2021). Another recent meta-analysis, but focused on the (partial) distance learning environment, also shows positive effects of motivation on academic achievement across the world. Indeed, highly motivated students could be more (cognitively) engaged in ((partial) distance) learning (Yu, 2021).

Third, academic self-efficacy can be described as an individual's conviction to successfully attain the desired academic goals (Bandura et al., 1999). This characteristic is positively associated with academic achievement across different educational stages, according to a systematic review of meta-analyses (Schneider & Preckel, 2017) and a more recent meta-analysis (Talsma et al., 2018). Meta-analytic evidence in a (partial) distance learning context shows that self-efficacy positively influences academic achievement as this characteristic could also greatly impact ((partial) distance) learning engagement (Yu, 2021), persistence etc. (Talsma et al., 2018). However, self-efficacy can also negatively impact academic achievement (Vancouver & Kendall, 2006), emphasizing the importance of distinguishing between the effort and comprehension dimensions of self-efficacy (Fonteyne et al., 2017). The effort dimension pertains to confidence in the ability to exert effort towards achieving academic goals and is positively associated with academic achievement. Conversely, the comprehension dimension relates to confidence in the ability to grasp course content. Overconfidence in the ability regarding this latter comprehension dimension can in fact reduce endeavor and subsequently result in diminished academic achievement (Fonteyne et al., 2017; Vancouver & Kendall, 2006).

At last, test anxiety can be defined as fear of or worry about negative evaluation (von der Embse et al., 2018) and shows a negative relation with academic achievement, as stated in a meta-analysis including compulsory and higher education studies (von der Embse et al., 2018). Some researchers find this negative influence of test anxiety on academic achievement as well when controlling for cognitive ability (Thomas et al., 2017). Further, the negative impact of test anxiety seems to be greater in an online proctored setting (Woldeab & Brothen, 2019). Other studies address that students with high test anxiety, in contrast, benefit more from online exams. However, these studies used an unproctored online setting (Stowell & Bennett, 2010) or online exams in a secure computer laboratory (Cassady & Gridley, 2005).

Interesting findings are found regarding the main effects of (non-)cognitive characteristics on academic achievement in the (partial) distance learning context. However, researchers rarely investigated the (non-)cognitive characteristics' moderating influences on the relationship between (partial) distance versus traditional learning and academic achievement. In COVID-19 studies additionally, (non-)cognitive factors' effects on academic achievement are currently understudied. Simultaneously, belief in the pandemic's different impact on students is omnipresent (Iterbeke & De Witte, 2021), which emphasizes the relevance of including student characteristics in pandemic research. The present study addresses this research gap by investigating interactions between the pandemic and (non-)cognitive characteristics in its effect on academic achievement in higher education.

Method

Participants

For the present study, we used secondary data from first-year university students of a large European university with eleven faculties and 42 bachelor's programs, ranked in the top 75 of the Academic Ranking of World Universities (formerly Shanghai Ranking, see https://www.shanghairanking.com/rankings/arwu/2022). The FPPW Ethics Committee at Ghent University provided favorable advice for the project (application number 2016/82). The university is characterized by an open access system[2] with strictly stratified study programs; full-time first-year students do have an identical curriculum within a study program. Only students who enrolled in an open access higher education study program for the first time in the academic years 2015-2016 to 2020-2021 and participated in the longitudinal university-wide study orientation project at the start of their first year in higher education (Fonteyne, 2017; Fonteyne et al., 2017) were considered. A wide range of student characteristics were selected and assessed through this platform, and linked to the first-year university students' academic achievement for the present study. This resulted in data from $N = 24{,}404$ (58% female, 23% low SES group) over a six-year period and across the 40 open access bachelor's programs. For more detail, see Appendix 2B, Tables B1 and B2.

Measures

Academic Achievement

In Belgium, an academic year in higher education is split into two semesters, each ending in a first-chance exam period. For each course, students receive a score from 0 to 20,

---

[2] With only one exemption for the study programs Medicine, Dentistry, and Performing and Visual Arts. Students have to pass an entrance exam to follow these programs. For other programs, secondary education qualifications suffice.

with a score of 10 necessary to pass. After the summer break, it is possible for students to retake an exam in the second-chance exam period if they failed on their first attempt. Furthermore, a number of ECTS credits (European Credit Transfer and Accumulation System credits) (European Commission, 2015) is linked to every course. Students in the model trajectory can take on and obtain a total of 60 ECTS credits per academic year. The distribution of these credits among the semesters depends on the study program, but is approximately balanced.

For the present study, academic achievement was operationalized by using the final study success scores, thus including the results of the second-chance exam period (August – September). This dependent variable, *study success*, shows the ratio of a student's obtained amount of ECTS credits over a student's subscribed amount of ECTS credits. Study success was scaled from 0 (failed all enrolled courses) to 100 (passed all enrolled courses).

Cohort

The cohort variable was used as a measure for the COVID-19 pandemic. In general, all Flemish universities switched to (partial) distance education since March 2020. Key to our between-subjects study design is that the cohorts 2015-2016 till 2018-2019 relate to the students who started their first year of higher education in a non-pandemic academic year. Cohort 2019-2020 consists of the first-year students who experienced one-third[3] of a pandemic academic year (see also Appendix 2A, Figures A1 and A2), while cohort 2020-2021 experienced their entire first year of higher education in full pandemic (see also Appendix 2A, Figures A1, A3 and A4).

[3] Note that two of the three exam periods (i.e., the first-chance exam period in the second semester and the second-chance exam period) did take place during the pandemic.

Student Characteristics

The longitudinal university-wide study orientation test battery (see also *Participants)* was used to measure the student characteristics. For an overview, please see Table 1.

Table 1

Overview Student Characteristics

Variable	Cat.	Values	Survey (example item)	n_{items}	α	*M* (*SD*)
Sex[1]	B	Male (0) Female (1)				
SES	B	High SES (0) Low SES (1)[2]				
Education Type SE[3]	C	General (0) Technical (1)				
Hours Maths SE	C					5.1 (1.8)
Vocabulary (/20)	C		LexTALE[4] ("Is this an existing Dutch word or not?")	60	.72	17.6 (1.7)
Self-Control (/20)	NC		Brief Self-Control Scale[5] ("I am able to work effectively toward long-term goals")	13	.75	13.0 (1.9)
Motivation:	NC		Academic Self-Regulation Questionnaire[6]			
Autonomous (/20)			("I study because I want to learn new things")	8	.86	15.0 (2.4)
Controlled (/20)			("I study because I am supposed to do so")	8	.87	8.3 (3.2)
Academic Self-Efficacy:	NC		College Academic Self-Efficacy Scale[7]			
Effort (/20)			("Attending class regularly")	8	.76	15.2 (1.9)
Comprehension (/20)			("Understanding most ideas you read in texts")	14	.81	14.8 (1.7)
Test Anxiety (/20)	NC		Cognitive Test Anxiety Scale Revised[8] ("I am not good at taking exams")	25	.92	10.0 (2.5)

Note. B = Background, C = Cognitive, NC = Non-Cognitive. SE = Secondary Education. [1]As stated on the passport. [2] Low-educated background (i.e., neither parent has completed secondary education) and/or receiving a scholarship. [3]General secondary education prepares students for higher education, while technical secondary education also prepares for professional careers and thus serves a dual purpose. [4]Lemhöfer and Broersma (2012). [5]Tangney and colleagues (2004). [6]Vansteenkiste and colleagues (2009). [7]Fonteyne and colleagues (2014) adapted from Owen and Froman (1988). The items are preceded by: "To what extent do you believe you are capable of performing each of the following tasks?" [8]Cassady and Finch (2015). Test Anxiety was measured through a 4-point Likert-scale, whereas a 5-point Likert-scale was used for the other non-cognitive variables to indicate the degree of agreement with the items.

Analyses

First, we examined the relationship between the pandemic and academic achievement in higher education, through a multilevel analysis using linear mixed-effects modeling. The cohort variable concerned the fixed factor and higher education study program the random factor. We calculated the conditional R-Squared (R_c^2) and marginal R-Squared (R_m^2) of the linear mixed model (Nakagawa & Schielzeth, 2012).[4] Further, Bonferroni-adjusted pairwise comparisons were performed to investigate the multiple comparisons of the estimated marginal means of study success (i.e., controlled for higher education study program as a random factor) between the cohorts. Because the cohorts have varying sample sizes, we used Hedges' *g* to calculate the effect sizes for the pairwise comparisons, which is adjusted based on the relative sample sizes. We applied the following rule of thumb: *g* = 0.10 (very small), *g* = 0.20 (small), *g* = 0.50 (medium), *g* = 0.80 (large), *g* = 1.20 (very large) and *g* = 2.00 (huge) (Marfo & Okyere, 2019).

Second, we implemented a stepwise selection using the AIC (Akaike's Information Criterion) procedure to identify the best predicting model for academic achievement (Burnham & Anderson, 2004). This stepwise selection method considers all possible models with all available predictors. Ultimately, the model with the lowest AIC is identified, ensuring the optimal balance between model complexity and goodness of fit. As such, information loss is minimalized, and overfitting is avoided by sanctioning excessive use of predictors. Unlike traditional stepwise regression, this AIC-driven approach avoids reliance on statistical tests for the model selection criterion and is independent of the order in which variables are introduced, evaluating every conceivable model with the potential predictors (Burnham & Anderson, 2004). This method was particularly useful given our model's requirement to

[4] Conditional R-Squared reflects the proportion of the total variance in the multilevel analysis that is explained by both fixed and random effects. Marginal R-Squared represents, on the other hand, the proportion of the variance explained solely by the fixed effects.

handle a considerable number of predictors, including cohort, the background and (non-)cognitive student characteristics and their interactions with cohort.

Third, a multilevel analysis on the resulting most optimal model to predict study success was conducted by using linear mixed-effects modeling. This time, the cohort variable and the other included background/(non-)cognitive student characteristics and interactions with cohort concerned the fixed factors and higher education study program the random factor. Prior to this analysis, we checked for multicollinearity by producing the Variance Inflation Factor (VIF) values for each of the independent variables. VIF values below ten are generally acceptable, but values above five can indicate significant multicollinearity. Therefore, maintaining VIF < 5 is recommended to ensure reliable results (Marcoulides & Raykov, 2019). We also calculated R_c^2 and R_m^2 of the linear mixed model, and both R_m^2 and unique R_m^2 values of the fixed effects. The latter effect size measure concerns the differences between the R_m^2 of the full model and the R_m^2 of the model without a specific fixed factor. See also Footnote 4.

Fourth, the significant interactions in this model according to the multilevel analysis were examined in more detail through Bonferroni-adjusted pairwise comparisons. These comparisons rely on the estimated marginal means, which are the means extracted from our most optimal statistical model and thus controlled for higher education study program as a random factor and for the other appearing predictors in the model. Hedges' g was used for the effect size calculations of these pairwise comparisons (Marfo & Okyere, 2019).

Results

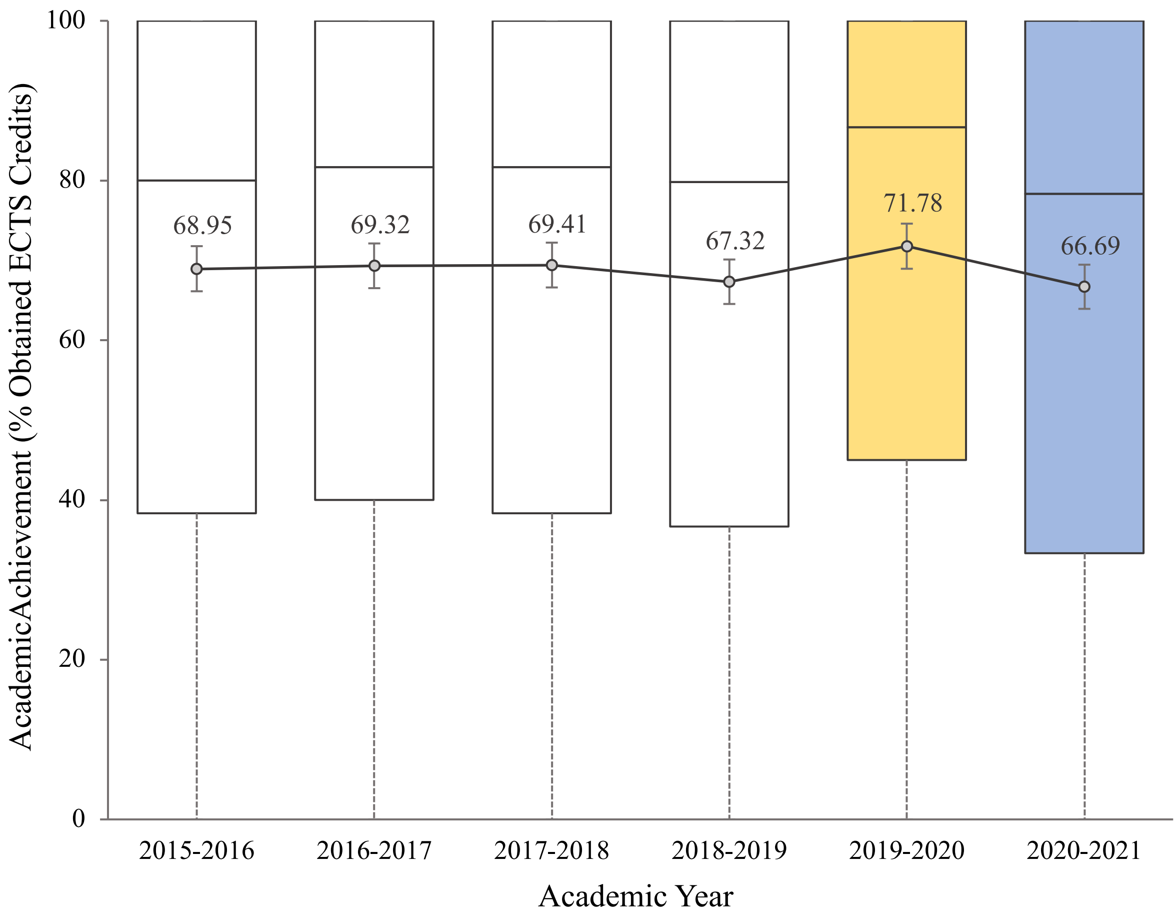
First, we investigated the effect of the pandemic on academic achievement (i.e., study success) in higher education, controlled for higher education study program as a random factor. Study success is expressed as the ratio between a student's obtained amount of ECTS credits over a student's subscribed amount of ECTS credits (%). In what follows, study success should always be interpreted as the estimated marginal mean study success (i.e., controlled for higher education study program as a random factor and for other appearing predictors in the model). The linear mixed-effects model shows a significant difference between study success of the cohorts (F(5, 24,377) = 11.04, p < .001, R^2_m = .002, R^2_c = .05). See also Figure 1. The Bonferroni-adjusted pairwise comparisons indicate that study success of the one-third pandemic cohort 2019-2020 (M = 71.78, SE = 1.41) is significantly higher in comparison with study success of the four pre-pandemic cohorts (difference range [2.36, 4.45], p's ≤ .049, g's ≤ 0.13). Further, study success of the full pandemic cohort 2020-2021 (M = 66.69, SE = 1.37[5]) is significantly lower than the pre-pandemic cohorts 2015-2016, 2016-2017 and 2017-2018 (difference range [2.26, 2.72], p's ≤ .041, g's ≤ 0.08). Additionally, cohort 2020-2021 shows lower study success than the year before (p < .001, g = 0.15). The differences in study success between two pre-pandemic cohorts are non-significant.[6] For the descriptives, multilevel analysis results and pairwise comparisons' extensive results, see Appendix 2B, Tables B3 to B4.

[5] Note that we find the largest variance for the full pandemic cohort 2020-2021.

[6] Multilevel analyses by using linear mixed-effects modeling and Bonferroni-adjusted pairwise comparisons between cohorts regarding the included (non-)cognitive variables show that minor fluctuations are found when controlled for higher education study program as a random factor, distributed across the cohorts. See also Appendix 2B, Tables B14 to B16.

Figure 1

Academic Achievement per Cohort



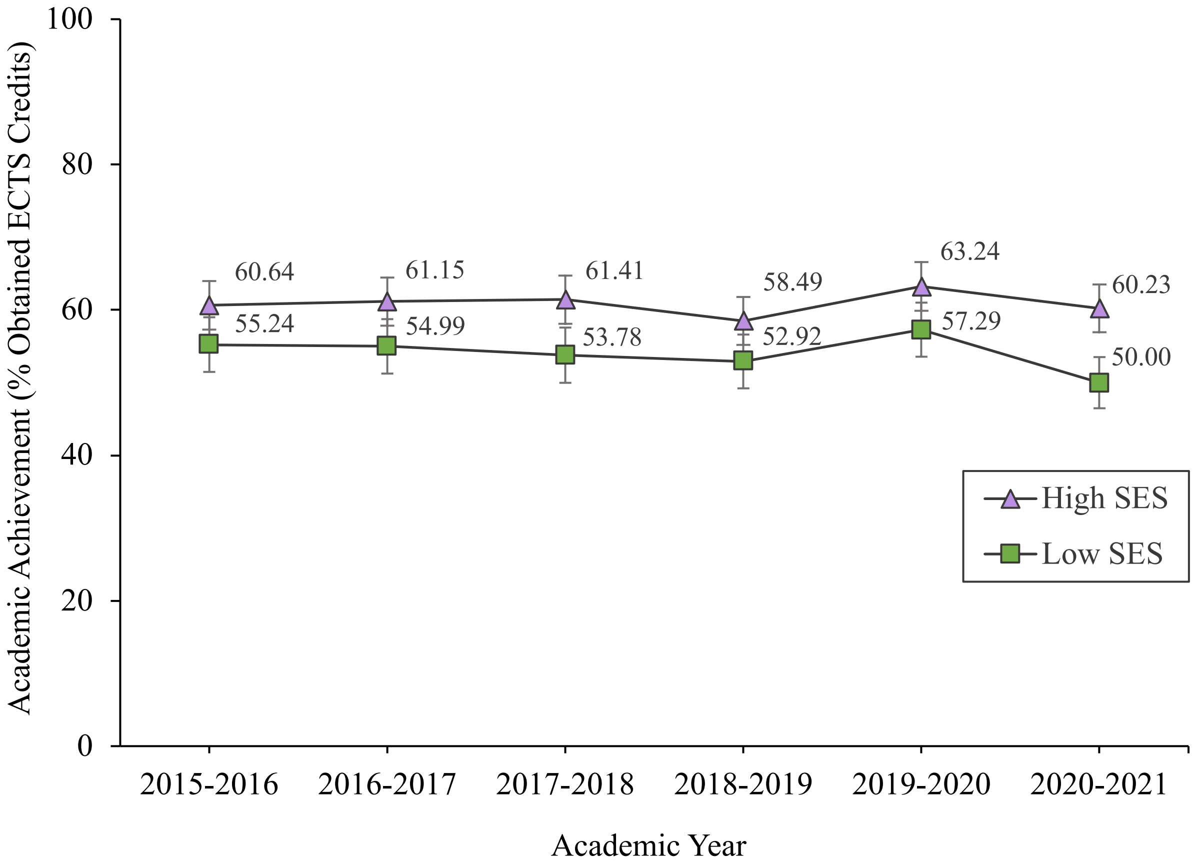
Note. Students' study success after the second-chance exam period (%) from cohorts 2015-2016 to 2020-2021. The white boxplots show the data distributions of the pre-pandemic cohorts, while the yellow and blue boxplots represent the cohorts of students who experienced a one-third and a full pandemic year, respectively. The estimated marginal means (i.e., controlled for higher education study program as a random factor) are represented by the grey dots and the grey error bars show the 95% confidence intervals of these means. See also Appendix 2B*,* Tables B3 and B4.

Second, we examined how the impact of the pandemic (i.e., the different cohorts) on academic achievement interacts with background and (non-)cognitive student characteristics. For the correlation matrix, we refer to Appendix 2B, Table B5. An AIC procedure (see *Analyses)* on a pool of predetermined possible predictors and interactions selected an optimal regression model for the prediction of study success. This final model contains (1) a set of predictor main effects including cohort, sex, SES, education type secondary education, hours of mathematics secondary education, vocabulary, self-control, self-efficacy (effort), self-efficacy (comprehension), test anxiety, autonomous motivation and controlled motivation, and (2) the interaction between the cohort variable and the student background variable SES. For a full overview, see Appendix 2B, Table B6. No VIF-value above 2 is present, indicating no multicollinearity issues.

To continue, a multilevel analysis through a linear mixed-effects model on this most optimal predictive model for study success was performed (R_c^2 = .23, R_m^2 = .17), which shows a significant interaction effect between cohort and SES (F(5, 24,348) = 2.92, p = .012). See also Figure 2. For the detailed multilevel analysis output, we refer to Appendix 2B, Tables B7 and B8. The estimates for the interaction between cohort and SES are shown in Appendix 2B, Table B9. Bonferroni-adjusted pairwise comparisons indicate that the study success of students of the low SES group is significantly lower than for students of the high SES group. This is the case for the four pre-pandemic cohorts (difference range [5.40, 7.63], p's $\leq$ .010, g's $\leq$ 0.24), and a similar effect is found for the one-third pandemic cohort 2019-2020 ($|\Delta EMM|$ = 5.95, p < .001, g = 0.19). The full pandemic cohort 2020-2021 shows a larger SES effect ($|\Delta EMM|$ = 10.23, p < .001, g = 0.32). Please see Appendix 2B, Table B10 for an overview.

Figure 2

Interaction Between Cohort and Socio-Economic Status on Academic Achievement



Note. Differences in high and low SES students' study success after the second-chance exam period (%) from cohorts 2015-2016 to 2020-2021, based on the estimated marginal means from our most optimal predictive model. Students from the cohorts 2015-2016 to 2018-2019, cohort 2019-2020 and cohort 2020-2021 experienced a 'normal', a one-third and a full pandemic academic year, respectively. The estimated marginal mean values (i.e., controlled for higher education study program as a random factor and for the other appearing predictors in the model) are represented by the markers and the error bars show the 95% confidence intervals of these means. See also Appendix 2B, Tables B9 and B10.

Discussion

To guarantee the continuity of learning, the COVID-19 pandemic required educational institutions to make an unprecedented and rapid shift to (partial) distance learning (Moscoviz

& Evans, 2022; OECD, 2021). As a cost of these educational institution closures, potential learning losses and reinforcement of pre-existing educational inequality were assumed and also observed (Azevedo et al., 2022; Betthäuser et al., 2023). To our knowledge, however, no study investigated possible learning losses in full pandemic year 2020-2021 compared to (a) pre-pandemic year(s) for higher education. Also, including both background as well as (non-)cognitive student characteristics as possible moderators in the relationship between the pandemic and academic achievement is often overlooked (Betthäuser et al., 2023; Iterbeke & De Witte, 2021). Consequently, the present empirical study used data from a running longitudinal project to investigate the pandemic's effect on academic achievement in higher education from a macro-level perspective. Additionally, we examined the interactions between the pandemic and a range of background and (non-)cognitive student characteristics that are related to academic achievement (Schneider & Preckel, 2017; Voyer & Voyer, 2014). Therefore, we utilized a large sample of more than 24,000 students of a Flemish, top-100 university, across six years. The final two cohorts experienced a one-third (2019-2020) and a full (2020-2021) pandemic year.

A surprising main finding of our study is the higher mean academic achievement of students from the one-third pandemic cohort 2019-2020 compared with the four pre-pandemic cohorts. Translated to ECTS credits, these differences correspond to one to three obtained ECTS credits[7] (to put this in perspective: three ECTS credits is equivalent to a small one-semester course). For the sake of argument, we also compared the mean academic achievement of this cohort with the four pre-pandemic year cohorts as a whole, controlled for higher education study program as a random factor. Students from cohort 2019-2020 show

[7] We calculated the amount of obtained ECTS credits by multiplying the percentage of academic achievement by the amount of subscribed ECTS credits. The number of subscribed ECTS credits is fixated on 60, as this is the standard program in the EU Bachelor (ISCED level 6) for first-year HE students. One ECTS-credit is equivalent to a mean of 30 study hours in order to successfully complete a course.

higher academic achievement than students from cohort 2015-2016 to 2018-2019, with a two obtained ECTS credits difference.[8] The increase in academic achievement for cohort 2019-2020 differs from what we expected and observed in compulsory education (Betthäuser et al., 2023), but the effect sizes are small. From a macro-level perspective, the crisis-response migration towards (partial) distance learning and the simultaneous other pandemic restrictions and health issues in 2019-2020 do not seem to negatively affect academic achievement in higher education. Some researchers address adaptation of learning content and exams, provision of alternatives for tasks and practical sessions, and/or grading differences as possible reasons for improved academic achievement during the pandemic (e.g., Bird et al., 2022; Rodríguez-Planas, 2022).[9] These factors, in turn, could have been beneficial in terms of additional learning time for the students, which is also voiced as a possible contributor to increased academic achievement in a blended learning environment (Means et al., 2013; Vo et al., 2017). In compulsory education, however, the amount of learning time halved from 7.50 hours to 3.70 hours per day during the first school closures (Werner & Woessmann, 2021). Students and parents appear to invest more time in learning activities when diversified educational inputs are provided (e.g., live contact hours) (Bansak & Starr, 2021). In March 2021, one year later, students are found to spend 4.60 hours per day on school-related activities (Werner & Woessmann, 2021). Further, some researchers conclude that a general

[8] We are aware of the difference in sample size between these groups. See also Appendix 2B, Tables B11 and B12.

[9] For the present study, possible indications of altered requirements during the pandemic in 2019-2020 can be found in study progress analyses presented for the Education Council (personal communication, March 17, 2022). These analyses show that by the end of 2019-2020, within the group of re-registrants (i.e., students who did not fully pass in their first year) the results are remarkably higher than in previous cohorts. Moreover, at the end of 2019-2020, fewer first-year students receive a binding condition (i.e., study success lower than 50% at the end of the first bachelor's year), but this group does score lower in 2020-2021 compared with previous cohorts. Similarly, re-registrants without a binding condition (i.e., more than 50% but less than 100% study success at the end of the first bachelor's year) show a decrease in academic achievement in 2020-2021 in comparison with cohort 2019-2020.

change in the autonomous learning process of students can be responsible for the improved academic achievement during the pandemic in higher education. In fact, the pandemic has caused a new scenario for the students. No previous similar experience may have contributed to more consistent course attendance by students and more continuous monitoring of their learning process (De Paolo et al., 2022).

On the other hand, the full pandemic cohort 2020-2021 shows the lowest mean academic achievement contrasted to the pre-pandemic cohorts 2015-2016 to 2017-2018. The differences are small and amount to approximately one to two obtained ECTS credits. The effect of the pandemic is thus limited, but important to point out. In fact, during the pre-pandemic years we observe normal fluctuations, while academic achievement differences are found between the full pandemic cohort 2020-2021 and certain pre-pandemic years. Similar to cohort 2019-2020, we also examined the mean academic achievement of cohort 2020-2021 with the four pre-pandemic year cohorts as a whole, controlled for higher education study program as a random factor. Students from cohort 2020-2021 show lower mean academic achievement than students from cohorts 2015-2016 to 2018-2019 (i.e., difference of one obtained ECTS-credit) (see Footnote 8). The many side effects of the pandemic seem not to have compensated for the favorable learning effects of (partial) distance learning in pre-pandemic times (e.g., Vo et al., 2017). The findings are partly in line with our expectation to observe learning losses, as perceived in compulsory education (Donnelly & Patrinos, 2021; Moscoviz & Evans, 2022). However, the effect sizes are rather small and not comparable to the months of cognitive delay demonstrated for younger children (i.e., PISA scores) (OECD, 2023). From a developmental perspective, this implication is not that surprising. Indeed, assumptions are made that younger students seem to rely more on cognitive scaffolding during instruction and their development of self-regulated learning skills might not yet be sufficient. Also, their vulnerability towards pandemic related stress might be higher than among older students. Consequently, the pandemic can hit them harder in their learning compared with the university students tested here (Tomasik et al., 2021).

When specifically comparing students from cohort 2019-2020 and cohort 2020-2021, the first cohort obtains about three ECTS credits more than students from the second cohort. Several factors could contribute to the accumulation of learning loss in the first full pandemic year 2020-2021 relative to the partial pandemic year 2019-2020. A first possible explanation is that students of cohort 2019-2020 experienced a 'normal' first two thirds of their first higher education year and thus went through their first exam period when there was no pandemic yet. In this way, they had already become acquainted with the functioning of higher education. Second, these students had been able to be more socially integrated in the academic environment what makes them better positioned to improve their academic achievement (Kassarnig et al., 2018; Rayle & Chung, 2007). These experiences do not apply to students from cohort 2020-2021, as they started their first year of higher education in full pandemic. Moreover, these students from cohort 2020-2021 also had to complete their last year of secondary education during the pandemic, where large learning losses are observed (Betthäuser et al., 2023; Donnelly & Patrinos, 2021; Moscoviz & Evans, 2022).

An additional interesting finding concerns the largest variance in academic achievement found for the full pandemic cohort 2020-2021. Some compulsory education studies also address the increase in heterogeneity of academic achievement during the pandemic (Tomasik et al., 2021). For example, the SES of students might explain the increasing heterogeneity in a full pandemic year, as low SES students experience little to no access to and less support in (partial) distance learning than high SES students (Kintu et al., 2017; Pokhrel & Chhetri, 2021). For more details, see Appendix 2B, Table B13. In what follows, we elaborate on the possible moderating role of background and (non-)cognitive student characteristics in the relationship between the pandemic and academic achievement.

Due to the unpredictability of the pandemic outbreak, the choice of these student characteristics was contingent on the accessibility of data from the running longitudinal project (see *Method*).

The present study shows that SES is a moderating factor between cohort and academic achievement in the most optimal predictive model, although the impact is rather small. As assumed and in line with (pre-)pandemic research (Betthäuser et al., 2023; Çiftçi & Cin, 2017), we find lower academic achievement for low SES students compared with high SES students. This academic achievement gap between the low SES and high SES students seems to be smaller for cohort 2019-2020 than for cohort 2020-2021, with a difference of four obtained ECTS credits in cohort 2019-2020. A possible explanation lies in the fact that in 2019-2020 the students first experienced a two-thirds 'normal' academic year, followed by the sudden shift to (partial) distance learning during the academic year. This quick crisis-response might have resulted in adaptions of learning content and assessment (Gonzalez et al., 2020; Iglesias-Pradas et al., 2021) to make the sudden and anything but easy situation for the students first and foremost more bearable. This approach could have been advantageous in terms of learning time and consequently academic achievement for both low and high SES students.

For the full pandemic cohort 2020-2021, on the contrary and as hypothesized, the largest academic achievement gap is found, with a difference of six obtained ECTS credits. These students had not yet experienced the normal course of events in higher education, making the accessibility and affordability to e-learning infrastructures (Azevedo et al., 2022; OECD, 2021) and a supportive environment (Hammerstein et al., 2021) definitely important. For low SES students moreover, obtaining these crucial factors is even more challenging (Azevedo et al., 2022; OECD, 2021), and they are also more likely to suffer from COVID-19 infection and mental distress (Anderson et al., 2020).

Interestingly, our findings also indicate that, besides SES, the pandemic does not interact with the many other student characteristics included in the present study. In line with previous COVID-19 research in higher education (El Said, 2021; Orlov et al., 2021), sex does not moderate the relationship between the pandemic and academic achievement. Furthermore, no interaction effect is found for education type (i.e., followed track) in secondary education, hours of mathematics in secondary education, vocabulary level, self-control, self-efficacy (effort), self-efficacy (comprehension), test anxiety, autonomous and controlled motivation. Note that despite the absence of such interactions, we do observe main effects on academic achievement for each and every of these characteristics in our most optimal predictive model. These student characteristics are thus important for academic achievement but do not affect the, from a macro-level perspective, influence of the pandemic learning situation.

When controlling for the other variables, first, we replicate better academic achievement for females (Schneider & Preckel, 2017). Second, hours of mathematics in secondary education and language proficiency positively influence academic achievement and following general secondary education (versus technical secondary education) is associated with higher academic achievement, which confirms the findings of existing literature (Ashford et al., 2016; Heeren et al., 2021). Third, the positive main effects of self-control, autonomous motivation, self-efficacy (effort) and negative main effect of test anxiety are also in line with previous research regarding the influence of these factors on academic achievement, both in a general (Howard et al., 2021; Talsma et al., 2018) and in a (partial) distance learning environment (Yu, 2021; Zhu et al., 2016). Fourth, in contrast to the effort dimension of self-efficacy, self-efficacy (comprehension) negatively impacts academic achievement. This finding is consistent with research that indicates that the comprehension dimension of self-efficacy can result in reduced endeavor and consequently in decreased academic achievement, due to one's overconfidence in comprehension abilities (Fonteyne et al., 2017; Vancouver & Kendall, 2006). Such evidence reinforces the necessity of addressing both self-efficacy dimensions distinctly to fully understand their implications on academic achievement. Finally, whereas previous research does not find a positive effect of controlled motivation on academic achievement (Howard et al., 2021), we do notice a main effect when controlling for the other included variables.

To our knowledge, investigating the potential influence of student characteristics (other than SES and sex) on the relationship between the pandemic and academic achievement in higher education is missing in published COVID-19 studies. We find significant main effects of these included student characteristics in combination with null interaction effects. These results, observed in a large sample with lots of statistical power, confirm that these characteristics are not moderators of the pandemic's effect on academic achievement, unlike SES.

Strengths, Limitations and Future Research Suggestions

The present study contributes to the currently scarce literature on the impact of the COVID-19 pandemic on academic achievement in higher education, which we approached from a macro-level perspective. The crisis-situation instigated a sudden and compelled shift in (inter)national educational organization toward (partial) distance learning, simultaneous with other (e.g., social) restrictions and health issues. Although the data analyzed in this study are from a single university, eleven faculties and 40 bachelor's programs were included and controlled for. As a result of our prospective study over a time period of six consecutive years, we can make statements about academic achievement in terms of study success (i.e., meeting minimum requirements) of different pre-pandemic cohorts compared with the one-third pandemic cohort 2019-2020 and the full pandemic cohort 2020-2021. Given the university's open access system (see also *Method*), the observed small impacts of the pandemic on academic achievement in higher education are even more striking.

Important to mention is the possibility of assessment biases, although the university does have standards, practices and testing systems in place. Educators had to make strategic choices that could have involved some degree of lenience at the onset of the pandemic in March 2020, followed by a potential greater emphasis on basic competences in the assessment periods. This approach might have resulted in the improved academic achievement in the one-third pandemic cohort 2019-2020. However, this comment also applies to compulsory education and the full pandemic cohort 2020-2021, in which academic achievement does deteriorate (to a limited extent for higher education).

Furthermore, including actual grades (e.g., GPA) as a measure of academic achievement was not possible here due to privacy reasons and ethical clearance, but is recommended for follow-up research. We do acknowledge that GPA is a well-established measure for assessing academic achievement (Richardson et al., 2012). However, GPA does not necessarily reflect the extent to which a student successfully completes their academic year. Students with similar GPAs may differ in the consistency of their academic performance and the total credits ultimately obtained. Indeed, GPA is more sensitive than study success to individual grades across diverse courses. Additionally, study success is based on ECTS credits, a recognized standardized system within European higher education designed primarily to measure and compare academic achievement across higher education institutions (European Commission, 2015). Also, we are aware that the measured academic achievement does not result from standardized testing. For higher education studies however, the use of interuniversity standardized exams is more difficult than in secondary education, given the unavailability of such exams in higher education.

As the pandemic was still ongoing during academic year 2021-2022, future research (in compulsory and higher education) could focus on academic achievement and student characteristics by considering more different pandemic and pre-pandemic years. Including post-pandemic years will also provide added value, even more so through a longitudinal approach.

Conclusion

The goal of the present study was to provide a more profound understanding of the pandemic's effects on academic achievement in higher education from a macro-level perspective, with specific consideration to the potential moderating role of student characteristics. The observed learning losses in higher education are rather small and more limited compared with compulsory education, with minor differences over SES. In particular, awareness of the vulnerability among the low SES students remains extremely important and additional support is recommended.

References

Adedoyin, O. B., & Soykan, E. (2020). Covid-19 pandemic and online learning: the challenges and opportunities. *Interactive Learning Environments, 31*(2), 863-875. https://doi.org/10.1080/10494820.2020.1813180

Anderson, G., Frank, J. W., Naylor, C. D., Wodchis, W., & Feng, P. (2020). Using socioeconomics to counter health disparities arising from the covid-19 pandemic. *BMJ*, 1-4. https://doi.org/10.1136/bmj.m2149

Asarta, C. J., & Schmidt, J. R. (2017). Comparing student performance in blended and traditional courses: Does prior academic achievement matter? *The Internet and Higher Education*, *32*, 29–38. https://doi.org/10.1016/j.iheduc.2016.08.002

Ashford, S. N., Lanehart, R. E., Kersaint, G. K., Lee, R. S., & Kromrey, J. D. (2016). STEM Pathways: Examining Persistence in Rigorous Math and Science Course Taking. *Journal of Science Education and Technology*, *25*(6), 961–975. https://doi.org/10.1007/s10956-016-9654-0

Azevedo, J. P., Hasan, A., Goldemberg, D., Aroob, I. S., & Geven, K. (2020). *Simulating the Potential Impacts of COVID-19 School Closures on Schooling and Learning Outcomes: A Set of Global Estimates.* http://hdl.handle.net/10986/33945

Azevedo, P., Gutierrez, M., de Hoyos, R., & Saavedra, J. (2022). The Unequal Impacts of COVID-19 on Student Learning. In F. M. Reimers (Ed.), *Primary and Secondary Education During Covid-19. Disruptions to Educational Opportunity During a Pandemic* (pp. 421–459). https://doi.org/10.1007/978-3-030-81500-4_16

Bandura, A., Freeman, W. H., & Lightsey, R. (1999). Self-Efficacy: The Exercise of Control. *Journal of Cognitive Psychotherapy*, *13*(2), 158–166. https://doi.org/10.1891/0889-8391.13.2.158

Bansak, C., & Starr, M. (2021). Covid-19 shocks to education supply: how 200,000 U.S. households dealt with the sudden shift to distance learning. *Review of Economics of the Household*, *19*(1), 63–90. https://doi.org/10.1007/s11150-020-09540-9

Berger, N., & Archer, J. (2016). School socio-economic status and student socio-academic achievement goals in upper secondary contexts. *Social Psychology of Education*, *19*(1), 175–194. https://doi.org/10.1007/s11218-015-9324-8

Bernard, R. M., Borokhovski, E., Schmid, R. F., Tamim, R. M., & Abrami, P. C. (2014). A meta-analysis of blended learning and technology use in higher education: from the general to the applied. *Journal of Computing in Higher Education*, *26*(1), 87–122. https://doi.org/10.1007/s12528-013-9077-3

Betthäuser, B. A., Bach-Mortensen, A. M., & Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. *Nature Human Behaviour*, *7*(3), 375–385. https://doi.org/10.1038/s41562-022-01506-4

Bettinger, E. P., Fox, L., Loeb, S., & Taylor, E. S. (2017). Virtual Classrooms: How Online College Courses Affect Student Success. *American Economic Review*, *107*(9), 2855–2875. https://doi.org/10.1257/aer.20151193

Bird, K. A., Castleman, B. L., & Lohner, G. (2022). Negative Impacts From the Shift to Online Learning During the COVID-19 Crisis: Evidence From a Statewide Community College System. *AERA Open*, *8*, 1-16. https://doi.org/10.1177/23328584221081220

Burnham, K. P., & Anderson, D. R. (2004). *Model Selection and Multimodel Inference*. Springer New York. https://doi.org/10.1007/b97636

Cai, Z., Fan, X., & Du, J. (2017). Gender and attitudes toward technology use: A meta-analysis. *Computers & Education*, *105*, 1–13. https://doi.org/10.1016/j.compedu.2016.11.003

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

Cassady, J. C., & Gridley, B. E. (2005). The Effects of Online Formative and Summative Assessment on Test Anxiety and Performance. *The Journal of Technology, Learning, and Assessment*, *4*(1), 4–30. https://ejournals.bc.edu/index.php/jtla/article/view/1648

Çiftçi, Ş. K., & Cin, F. M. (2017). The Effect of Socioeconomic Status on Students' Achievement. In *The Factors Effecting Student Achievement* (pp. 171–181). Springer International Publishing. http://doi.org/10.1007/978-3-319-56083-0_10

De Paolo, M., Gioia, F., & Scoppa, V. (2022). *Online Teaching, Procrastination and Students' Achievement: Evidence from COVID-19 Induced Remote Learning*. IZA Institute of Labor Economics Discussion Paper Series No. 15031. https://docs.iza.org/dp15031.pdf

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

Donnelly, R., & Patrinos, H. A. (2021). Learning loss during Covid-19: An early systematic review. *PROSPECTS, 10*, 1-9. https://doi.org/10.1007/s11125-021-09582-6

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*(1), 373–399. https://doi.org/10.1146/annurev-psych-010418-103230

El Said, G. R. (2021). How Did the COVID-19 Pandemic Affect Higher Education Learning Experience? An Empirical Investigation of Learners' Academic Performance at a University in a Developing Country. *Advances in Human-Computer Interaction*, *2021*, 1–10. https://doi.org/10.1155/2021/6649524

Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, *118*(17), 1-7. https://doi.org/10.1073/pnas.2022376118

European Commission, D.-G. for E. Y. S. and C. (2015). *ECTS users' guide 2015*. https://data.europa.eu/doi/10.2766/87192

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]. Ghent University.

Fonteyne, L., Duyck, W., & de Fruyt, F. (2017). Program-specific prediction of academic achievement on the basis of cognitive and non-cognitive factors. *Learning and Individual Differences*, *56*, 34–48. https://doi.org/10.1016/j.lindif.2017.05.003

Gonzalez, T., de la Rubia, M. A., Hincz, K. P., Comas-Lopez, M., Subirats, L., Fort, S., & Sacha, G. M. (2020). Influence of COVID-19 confinement on students' performance in higher education. *PLOS ONE*, *15*(10), 1-23. https://doi.org/10.1371/journal.pone.0239490

Hammerstein, S., König, C., Dreisörner, T., & Frey, A. (2021). Effects of COVID-19-Related School Closures on Student Achievement-A Systematic Review. *Frontiers in Psychology*, *12*, 1-8. https://doi.org/10.3389/fpsyg.2021.746289

Heeren, J., Speelman, D., & de Wachter, L. (2021). A practical academic reading and vocabulary screening test as a predictor of achievement in first-year university students: implications for test purpose and use. *International Journal of Bilingual Education and Bilingualism*, *24*(10), 1458–1473. https://doi.org/10.1080/13670050.2019.1709411

Howard, J. L., Bureau, J., Guay, F., Chong, J. X. Y., & Ryan, R. M. (2021). Student Motivation and Associated Outcomes: A Meta-Analysis From Self-Determination Theory. *Perspectives on Psychological Science*, *16*(6), 1300–1323. https://doi.org/10.1177/1745691620966789

Iglesias-Pradas, S., Hernández-García, Á., Chaparro-Peláez, J., & Prieto, J. L. (2021). Emergency remote teaching and students' academic performance in higher education during the COVID-19 pandemic: A case study. *Computers in Human Behavior*, *119*, 1-18. https://doi.org/10.1016/j.chb.2021.106713

Iterbeke, K., & De Witte, K. (2021). Helpful or Harmful? The Role of Personality Traits in Student Experiences of the COVID-19 Crisis and School Closure. *Personality and Social Psychology Bulletin, 48*(11), 1614-1632. https://doi.org/10.1177/01461672211050515

Kassarnig, V., Mones, E., Bjerre-Nielsen, A., Sapiezynski, P., Dreyer Lassen, D., & Lehmann, S. (2018). Academic performance and behavioral patterns. *EPJ Data Science*, *7*(1), 1-16. https://doi.org/10.1140/epjds/s13688-018-0138-8

Kintu, M. J., Zhu, C., & Kagambe, E. (2017). Blended learning effectiveness: the relationship between student characteristics, design features and outcomes. *International Journal of*

Educational Technology in Higher Education, *14*(1), 1-20. https://doi.org/10.1186/s41239-017-0043-4

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Lichand, G., Doria, C. A., Leal-Neto, O., & Fernandes, J. P. C. (2022). Publisher Correction: The impacts of remote learning in secondary education during the pandemic in Brazil. *Nature Human Behaviour*, *6*(8), 1180–1180. https://doi.org/10.1038/s41562-022-01420-9

Liu, J., Peng, P., & Luo, L. (2020). The Relation Between Family Socioeconomic Status and Academic Achievement in China: A Meta-analysis. *Educational Psychology Review*, *32*(1), 49–76. https://doi.org/10.1007/s10648-019-09494-0

López-Pérez, M. V., Pérez-López, M. C., & Rodríguez-Ariza, L. (2011). Blended learning in higher education: Students' perceptions and their relation to outcomes. *Computers & Education*, *56*(3), 818–826. https://doi.org/10.1016/j.compedu.2010.10.023

Marcoulides, K. M., & Raykov, T. (2019). Evaluation of Variance Inflation Factors in Regression Models Using Latent Variable Modeling Methods. *Educational and Psychological Measurement*, *79*(5), 874–882. https://doi.org/10.1177/0013164418817803

Marfo, P., & Okyere, G. A. (2019). The accuracy of effect-size estimates under normals and contaminated normals in meta-analysis. *Heliyon*, *5*(6), 1-9. https://doi.org/10.1016/j.heliyon.2019.e01838

McQuaid, R. J., Cox, S. M. L., Ogunlana, A., & Jaworska, N. (2021). The burden of loneliness: Implications of the social determinants of health during COVID-19. *Psychiatry Research*, *296*, 1-7. https://doi.org/10.1016/j.psychres.2020.113648

Means, B., Toyama, Y., Murphy, R., & Baki, M. (2013). The Effectiveness of Online and Blended Learning: A Meta-Analysis of the Empirical Literature. *Teachers College Record*, *115*(3), 1–47. https://www.researchgate.net/publication/286792735_The_Effectiveness_of_Online_and_Blended_Learning_A_Meta-Analysis_of_the_Empirical_Literature

Moscoviz, L., & Evans, D. K. (2022). *Learning Loss and Student Dropouts during the COVID-19 Pandemic: A Review of the Evidence Two Years after Schools Shut Down*. https://www.cgdev.org/sites/default/files/learning-loss-and-student-dropouts-during-covid-19-pandemic-review-evidence-two-years.pdf

Nakagawa, S., & Schielzeth, H. (2012). A general and simple method for obtaining R^2 from generalized linear mixed-effects model. *Methods in Ecology and Evolution, 4*(2), 133-142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Nguyen, T. (2015). The Effectiveness of online learning: Beyond No Significant Difference and Future Horizons. *MERLOT Journal of Online Learning and Teaching*, *11*(2), 309-319. https://www.researchgate.net/publication/308171318_The_Effectiveness_of_Online_Learning_Beyond_No_Significant_Difference_and_Future_Horizons

OECD. (2021). *The State of Global Education - 18 Months into the Pandemic*. https://doi.org/10.1787/1a23bb23-en

OECD. (2023). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. https://doi.org/https://doi.org/10.1787/53f23881-en

Orlov, G., McKee, D., Berry, J., Boyle, A., DiCiccio, T., Ransom, T., Rees-Jones, A., & Stoye, J. (2021). Learning during the COVID-19 pandemic: It is not who you teach, but how you teach. *Economics Letters*, *202*, 1-4. https://doi.org/10.1016/j.econlet.2021.109812

Owen, S. v., & Froman, R. D. (1988). *Development of a College Academic Self-Efficacy Scale*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans, April 6-8.

Pei, L., & Wu, H. (2019). Does online learning work better than offline learning in undergraduate medical education? A systematic review and meta-analysis. *Medical Education Online*, *24*(1), 1-13. https://doi.org/10.1080/10872981.2019.1666538

Pierre, G., Sanchez Puerta, M. L., Valerio, A., & Rajadel, T. (2014). *STEP skills measurement surveys: innovative tools for assessing skills*. World Bank Group Working Paper 89729. https://openknowledge.worldbank.org/bitstream/handle/10986/19985/897290NWP0P132085290B00PUBLIC001421.pdf?sequence=1&isAllowed=y

Pokhrel, S., & Chhetri, R. (2021). A Literature Review on Impact of COVID-19 Pandemic on Teaching and Learning. *Higher Education for the Future*, *8*(1), 133–141. https://doi.org/10.1177/2347631120983481

Prowse, R., Sherratt, F., Abizaid, A., Gabrys, R. L., Hellemans, K. G. C., Patterson, Z. R., & McQuaid, R. J. (2021). Coping With the COVID-19 Pandemic: Examining Gender Differences in Stress and Mental Health Among University Students. *Frontiers in Psychiatry*, *12,* 1-11. https://doi.org/10.3389/fpsyt.2021.650759

Rayle, A. D., & Chung, K.-Y. (2007). Revisiting First-Year College Students' Mattering: Social Support, Academic Stress, and the Mattering Experience. *Journal of College Student Retention: Research, Theory & Practice*, *9*(1), 21–37. https://doi.org/10.2190/X126-5606-4G36-8132

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. https://doi.org/10.1037/a0026838

Rodríguez-Hernández, C. F., Cascallar, E., & Kyndt, E. (2020). Socio-economic status and academic performance in higher education: A systematic review. *Educational Research Review*, *29*, 1-24. https://doi.org/10.1016/j.edurev.2019.100305

Rodríguez-Planas, N. (2022). COVID-19, college academic performance, and the flexible grading policy: A longitudinal analysis. *Journal of Public Economics*, *207*, 1-11. https://doi.org/10.1016/j.jpubeco.2022.104606

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. https://doi.org/10.1016/j.intell.2015.09.002

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, *143*(6), 565–600. https://doi.org/10.1037/bul0000098

Schult, J., Mahler, N., Fauth, B., & Lindner, M. A. (2021). Did students learn less during the COVID-19 pandemic? Reading and math competencies before and after the first pandemic wave. *School Effectiveness and School Improvement, 33*(4), 544-563. https://doi.org/10.1080/09243453.2022.2061014

Spitzer, M. W. H., & Musslick, S. (2021). Academic performance of K-12 students in an online-learning environment for mathematics increased during the shutdown of schools in wake of the COVID-19 pandemic. *PLOS ONE*, *16*(8), 1-16. https://doi.org/10.1371/journal.pone.0255629

Stadler, M., Aust, M., Becker, N., Niepel, C., & Greiff, S. (2016). Choosing between what you want now and what you want most: Self-control explains academic achievement beyond cognitive ability. *Personality and Individual Differences*, *94*, 168–172. https://doi.org/10.1016/j.paid.2016.01.029

Stowell, J. R., & Bennett, D. (2010). Effects of Online Testing on Student Exam Performance and Test Anxiety. *Journal of Educational Computing Research*, *42*(2), 161–171. https://doi.org/10.2190/EC.42.2.b

Talsma, K., Schüz, B., Schwarzer, R., & Norris, K. (2018). I believe, therefore I achieve (and vice versa): A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Learning and Individual Differences*, *61*, 136–150. https://doi.org/10.1016/j.lindif.2017.11.015

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. https://doi.org/10.1111/j.0022-3506.2004.00263.x

Thomas, C. L., Cassady, J. C., & Heller, M. L. (2017). The influence of emotional intelligence, cognitive test anxiety, and coping strategies on undergraduate academic performance. *Learning and Individual Differences*, *55*, 40–48. https://doi.org/10.1016/j.lindif.2017.03.001

Tomasik, M. J., Helbling, L. A., & Moser, U. (2021). Educational gains of in-person vs. distance learning in primary and secondary schools: A natural experiment during the COVID-19 pandemic school closures in Switzerland. *International Journal of Psychology*, *56*(4), 566–576. http://doi.org/10.1002/ijop.12728https://doi.org/10.1002/ijop.12728

United Nations. (2020). *Education during COVID-19 and beyond*. https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2020/08/sg_policy_brief_covid-19_and_education_august_2020.pdf

Vancouver, J. B., & Kendall, L. N. (2006). When self-efficacy negatively relates to motivation and performance in a learning context. *Journal of Applied Psychology*, *91*(5), 1146–1153. https://doi.org/10.1037/0021-9010.91.5.1146

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. https://doi.org/10.1037/a0015083

Vo, H. M., Zhu, C., & Diep, N. A. (2017). The effect of blended learning on student performance at course-level in higher education: A meta-analysis. *Studies in Educational Evaluation*, *53*, 17–28. https://doi.org/10.1016/j.stueduc.2017.01.002

Vo, M. H., Zhu, C., & Diep, A. N. (2020). Students' performance in blended learning: disciplinary difference and instructional design factors. *Journal of Computers in Education*, *7*(4), 487–510. https://doi.org/10.1007/s40692-020-00164-7

von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders*, *227*, 483–493. https://doi.org/10.1016/j.jad.2017.11.048

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174–1204. https://doi.org/10.1037/a0036620

Werner, K., & Woessmann, L. (2021). *The Legacy of Covid-19 in Education.* IZA Institute of Labor Economics Discussion Paper Series No.9358. https://docs.iza.org/dp14796.pdf

Woldeab, D., & Brothen, T. (2019). 21st Century assessment: Online proctoring, test anxiety, and student performance. *International Journal of E-Learning & Distance Education, 34,* 1-10. http://www.ijede.ca/index.php/jde/article/view/1106

Yu, Z. (2021). A meta-analysis and bibliographic review of the effect of nine factors on online learning outcomes across the world. *Education and Information Technologies, 19,* 33-50. http://doi.org/10.1007/s10639-021-10720-y

Yu, Z., & Yu, L. (2021). A Meta-Analysis of Online Learning Outcomes and Their Gender Differences. *International Journal of Distance Education Technologies*, *19*(3), 33–50. https://doi.org/10.4018/IJDET.2021070103

Zhu, Y., Au, W., & Yates, G. (2016). University students' self-control and self-regulated learning in a blended course. *The Internet and Higher Education*, *30*, 54–62. https://doi.org/10.1016/j.iheduc.2016.04.001

3

Validating GULS: an Open Access Dutch Language Proficiency Test

Prepared for Submission to Behavior Research Methods, 26th of June 2024.

Mona Bassleer, Stijn Schelfhout, Lot Fonteyne, Mit Leuridan, Wouter Duyck & Nicolas Dirix.

Abstract

Previous post-entry language assessment (PELA) research in higher education shows that academic language proficiency contributes to academic achievement. PELAs are particularly valuable for higher education systems with minimal to no admission requirements to identify and support at-risk students. However, the availability of PELAs for languages like Dutch, beyond English, is limited. Moreover, existing Dutch PELAs and its construct validity evidence are not publicly accessible, and predictive validity analyses typically do not reach the program-specific level. Therefore, the present study introduced the Ghent University Language Screening (GULS), an easy-to-administer, free and publicly accessible Dutch PELA. More specifically, GULS evaluates reading comprehension of first-year students in higher education. First, we confirmed the construct validity of GULS at model and item level and its reliability using data from the two three-year periods 2017-2018 to 2019-2020 (N_1 = 12,527) and 2020-2021 to 2022-2023 (N_2 = 17,204). Second, we examined GULS's predictive validity for academic achievement (i.e., Grade Point Average and study success) on data over the same two periods across 16 bachelor's study programs (n_1 = 8,244; n_2 = 10,891), followed by predictive validity analyses for each study program across the combined six-year period. Results demonstrate that GULS is valid and reliable to assess Dutch language proficiency in terms of reading comprehension, especially for first-year higher education students requiring language support to ensure equal educational opportunities. As such, GULS functions as a predictor of first-year academic achievement. We discuss the possible application of GULS in future educational research and practice due to its accessibility.

Introduction

Today's 21st-century society requires adequate language proficiency from individuals for a full-fledged societal participation (Kennedy & Sundberg, 2020; OECD, 2019). In addition to everyday language proficiency, academic language proficiency is important for students entering higher education (Knoch & Elder, 2013; Read, 2016). This addition aligns with Hulstijn's (2015) concept of Higher Language Cognition (HLC), which stands in contrast to Basic Language Cognition (BLC). BLC pertains to the language proficiency required for everyday communication, whereas HLC as equivalent of academic language proficiency (hereafter: language proficiency) involves more advanced language skills at lexical, syntactical, and cognitive levels (Hulstijn, 2015). Additionally, reading comprehension, as a component of language proficiency, appears to be a challenging receptive skill predominantly relied upon by students during the early stages of higher education (De Wachter et al., 2013; Jansen et al., 2022; Van Houtven et al., 2010). The present study considers language proficiency in terms of reading comprehension, defined as understanding, using, evaluating, and reflecting (on) textual content (OECD, 2023). Following Kintsch's (2013) Construction-Integration model, an individual constructs meaning from and integrates this textual content with existing long-term knowledge during the reading comprehension process. Furthermore, this ability relies on coordinating lower-order skills like decoding with higher-order ones such as vocabulary knowledge use (Kintsch, 2013).

Language proficiency is known for playing a crucial yet not exclusive role in academic achievement (Elder, 2017; Heeren et al., 2021; Read, 2016). This ability is indeed used across different subjects and academic disciplines for content comprehension (Read, 2015). Various meta-analytic evidence shows a clear association between language proficiency of mainly non-native speakers and academic achievement (i.e., GPA), with average correlation coefficients ranging from $r = .18$ to $r = .23$ (Abunawas, 2014; Gagen, 2019; Ihlenfeldt & Rios, 2023; Wongtrirat, 2010). Nevertheless, language proficiency can challenge all (prospective) enrolling students in higher education (Elder, 2017; Wingate, 2015). In a recent meta-analysis, researchers specifically examined the reading comprehension components of more general admission assessments (e.g., SAT), showing a correlation coefficient of r = .29 (Clinton-Lisell et al., 2022). However, these instances of language proficiency assessment determine admission eligibility at English-medium higher education institutions.

Alternatively, higher education institutions with low or no admission requirements use post-entry language assessments (PELAs) to identify at-risk students and provide language support to enhance their academic journey. Indeed, the low-stakes environments of these institutions attract more diverse student populations enrolling in higher education (Elder, 2017; Knoch & Elder, 2013; Read, 2016). To the best of our knowledge, research on the predictive validity of PELAs for all enrolled higher education students on academic achievement is rather scarce (e.g., De Wachter et al., 2013; Heeren et al., 2021; van Dijk, 2015). Moreover, the availability of PELAs for languages other than English is currently very limited. In general, a distinct relationship is observed in (Dutch) PELA studies, with language proficiency accounting for a maximum of 10% of the variance in academic achievement (Elder, 2017; Heeren et al., 2021; Knoch & Elder, 2013). The present study therefore focuses on the validation of a Dutch PELA. Dutch is the official language of higher education in Flanders and the Netherlands, with over 24,300,000 native speakers (Eberhard et al., 2024). Governing bodies explicitly advocate for maintaining the Dutch language proficiency in education and society, especially amidst the increasing language diversity (Jansen et al., 2022). Moreover, Dutch holds merits in scientific research exemplified by its status as one of the most studied languages in psycholinguistics (Siegelman et al., 2022).

Researchers highlight the practice of internally developed PELAs, yet often without professional validation (Knoch & Elder, 2013). A recent review similarly underscores that over half of their included language assessment studies fail to address construct validity, and those that do often limit their analysis to the overall model-data fit (Min & Aryadoust, 2021). The test developed by Heeren and colleagues (2021) is a good example of a Dutch PELA to screen first-year university students in an open access[1] higher education system and identify those needing support. To our knowledge, however, both the detailed construct validation evidence as well as the test itself are not publicly accessible, compromising transparency, replicability, and reproducibility (Min & Aryadoust, 2021). Additionally, the predictive validity of this Dutch PELA was examined on first-year academic achievement in terms of credit completion rate, whereas the most common metric of academic achievement in educational research is Grade Point Average (GPA) (York et al., 2015). We include both GPA and study success (i.e., credit completion rate) as measures of academic achievement, recognizing the differing views on GPA's utility (Ihlenfeldt & Rios, 2023). Furthermore, assessing the predictive validity of language proficiency for academic achievement does not usually include separate analyses by study program. However, such methodological refinement can demonstrate the differential impact of language proficiency on academic achievement within distinct fields of study (Elder, 2017; Hauspie et al., 2024; Read, 2016).

To address the issues with and therefore advance existing PELA in higher education, the present study validates the Ghent University Language Screening (GULS), developed as a Dutch PELA with fully open access (i.e., easy to administer, free and publicly available). More specifically, GULS assesses reading comprehension of first-year higher education

[1] Students who successfully completed secondary education can enroll in higher education without admission requirements like an entrance exam, except for the study programs Medicine, Dentistry, and Performing and Visual Arts.

students. First, we examine and confirm GULS's construct validity on model and item level, and its reliability using extensive data from the academic years 2017-2018 to 2019-2020 and 2020-2021 to 2022-2023. Second, we evaluate the predictive validity of GULS for academic achievement (i.e., GPA and study success) on large prospective data from the same two three-year periods across 16 bachelor's study programs, culminating in predictive validity analyses for each study program spanning this extended aggregate six-year period. As such, the present study provides insights into the contribution of Dutch language proficiency to first-year academic achievement in various study programs, such as linguistic but also social science and STEM-oriented study programs. For these analyses, we also include control variables like socio-economic status (SES) and the mathematics category of the students' enrolled study program.[2] Practically, GULS can help inform (prospective) first-year higher education students about their language proficiency level, and guide remediation for those students at risk. Moreover, GULS's open access supports its potential use in future educational research.

Method

Participants

The present study used secondary data from first-year students at a large European university with eleven faculties and 42 bachelor's programs, consistently ranked in the top 75 worldwide according to the Academic Ranking of World Universities (formerly Shanghai Ranking, see https://www.shanghairanking.com/rankings/arwu/2022). The FPPW Ethics Committee at Ghent University granted approval. The university's open access system (see also Footnote 1) ensures a uniform curriculum for full-time first-year students within a study program. Our samples included students who entered higher education for the first time between 2017-2018 and 2022-2023 and participated in the university's study (re)orientation

---

[2] The present study differentiates between study programs that expect basic or advanced levels of mathematical proficiency. The advanced mathematics study programs are thus linked to more advanced mathematical curricula (see also *Measures*).

and remediation SIMON project. SIMON is a self-assessment tool that evaluates academic aptitude and vocational interests, guiding (prospective) students in their educational paths (Fonteyne, 2017; Fonteyne et al., 2017). GULS is a component of the wider SIMON test battery, administered to first-year students at the beginning of each academic year (i.e., October).

Construct Validity and Reliability

For the construct validity and reliability analyses, we used two samples. Only students with complete responses to the GULS, sex, and socio-economic status (SES) items (for the operationalizations, see *Measures*) in the SIMON tool were incorporated in both samples. Sample 1 comprised N_1 = 12,527 first-year higher education students from 2017-2018 to 2019-2020 (25% low SES; 40% male), and sample 2 included N_2 = 17,204 first-year higher education students from 2020-2021 to 2022-2023 (22% low SES; 41% male).

Predictive Validity

For the predictive validity analyses, we employed a subsample from each of the two original samples. Specifically, we included only the students enrolled in study programs for which we had at least 50 total responses for all six academic years to ensure sufficient statistical power. Here, total responses refer to complete responses on GULS and the control variables (see below), along with recorded GPA and study success scores after the second-chance exam period (i.e., August/September). As such, subsample 1 consisted of n_1 = 8,244 first-year higher education students from 2017-2018 to 2019-2020, and subsample 2 comprised n_2 = 10,891 first-year higher education students from 2020-2021 to 2022-2023, both across 16 study programs. Table 1 outlines these subsample sizes categorized by sex, SES, and the mathematics category of the students' enrolled study program. For the subsample sizes per study program, see Appendix 3A, Table A1.

Table 1

Subsample Sizes by Sex, SES, and Mathematics Category

	Subsample 1	Subsample 2	Total
Sex			
Male	3,048	4,208	7,256
Female	5,196	6,683	11,879
SES			
Low SES	2,085	2,363	4,448
High SES	6,159	8,528	14,687
Mathematics Category			
Basic	4,847	6,487	11,334
Advanced	3,397	4,404	7,801

Note. Subsamples 1 and 2 respectively encompass students from academic years 2017-2018 to 2019-2020 and 2020-2021 to 2022-2023. The advanced mathematics study programs are associated with more advanced mathematical curricula. Through the SIMON project (Fonteyne, 2017), students enrolling in basic mathematics study programs complete a less difficult math assessment compared to students enrolling in more advanced mathematics study programs. For more information on the predictive validity of this basic mathematics test, see Fonteyne and colleagues (2015).

Measures

Academic Achievement

On the one hand, academic achievement as the dependent variable in the present study was operationalized by using students' *Grade Point Average (GPA)*. At Ghent University, students have two first-chance exam periods, one per semester. Each course is graded from 0 to 20, with 10 as the passing mark. After summer break, a second-chance exam period is provided for students who did not pass one or more courses on their first attempt. Every course is also assigned a number of ECTS credits (European Credit Transfer and Accumulation System credits) (European Commission, 2015), which reflect its weight and are included in the GPA calculation. In the present study, GPA was scored from 0 to 100.

On the other hand, *study success* was also used as an academic achievement measure. A first-year student in the standard curriculum can accumulate a maximum of 60 ECTS credits. Study success was defined as the ratio of students' obtained ECTS credits over their subscribed ones after the second-chance exam period and ranged between 0 and 100.

Language Proficiency

Language Proficiency was operationalized by using GULS, which constitutes the central focus of the present study. The aim of developing GULS was to integrate the test into the existing SIMON tool for (prospective) higher education students, focusing on study (re)orientation and remediation (Fonteyne, 2017; Fonteyne et al., 2017). Recognizing the need for a comprehensive language proficiency test beyond vocabulary, developers piloted the initial 45-item GULS with last-year secondary students in 2016-2017. A team of experts, including linguists and experimental and cognitive psychologists, developed and (re)evaluated the items. To optimize student testing time and based on preliminary analyses that showed promising outcomes regarding the total scores' descriptives, reliability, and item difficulty and discrimination parameters, the developers subsequently shortened GULS to 25 items. For further details about the preliminary analyses of the pilot study, refer to Appendix 3C. The 25-item GULS can be found in Appendix 3D. In the present study, we assessed the construct/predictive validity and reliability of the 25-item GULS using data from the academic years 2017-2018 to 2019-2020, resulting in a refined set of GULS items. This selected set of GULS items served as the basis for the subsequent construct/predictive validity and reliability analyses on the data from the academic years 2020-2021 to 2022-2023 (see also *Analyses*).

Control Variables

Sex distinguished between males (= 0) and females (= 1) as stated on the student's passport.

Socio-Economic Status (*SES*) considered a student's scholarship status and their parents' educational attainment. Students are categorized into the low SES group (= 0) versus the high SES group (= 1) if neither parent completed secondary education and/or if they receive a scholarship. Such an operationalization of SES is recognized as a relevant predictor of academic achievement in previous research (e.g., Fonteyne, 2017; Hauspie et al., 2024).

Mathematics Hours Secondary Education refers to the students' self-reported total number of hours mathematics instruction they received during their last year of secondary education. Similar to the personal background characteristic SES, this educational background characteristic is an established predictor of academic achievement (e.g., Fonteyne, 2017; Schelfhout, 2019).

Vocabulary Knowledge was measured through the LexTALE test (Lemhöfer & Broersma, 2012). Students were asked to indicate, for each of the 60 items (ω's > .70), whether the item represented an actual word or not.

Mathematics Category refers to the category in which the studied study programs were classified: basic (= 0) and advanced (= 1), with the latter featuring more advanced mathematical curricula. This control variable was specifically used in the analyses regarding GULS's predictive validity for academic achievement on the 2017-2018 to 2019-2020 and 2020-2021 to 2022-2023 data across study programs.

Mathematical Proficiency was measured using either the basic or advanced mathematics test within the SIMON tool (Fonteyne, 2017), depending on the student's enrolled study program (see also *Mathematics Category*). The basic mathematics test comprised 20 items (open and multiple-choice; ω's > .70), while the advanced one consisted of 25 multiple-choice items (ω's > .70) and was more challenging. For examples and more information on the predictive validity of the basic mathematics test, see Fonteyne (2017) and Fonteyne and colleagues (2015). This control variable was specifically used in the analyses regarding GULS's predictive validity for academic achievement on the 2017-2018 to 2022-2023 data per study program.

Analyses

Construct Validity and Reliability

First, we evaluated the construct validity and reliability of GULS on the data from 2017-2018 to 2019-2020 (i.e., sample 1), with 25 GULS items. A Univariate Variable Analysis (UVA) was conducted to assess local independence. This analysis identifies items with a weighted Topological Overlap (wTO) exceeding .25, suggesting potential interdependence and recommending removal of one item to mitigate redundancy (Christensen et al., 2023). Building upon literature on language proficiency, we proceeded with a Confirmatory Factor Analysis (CFA) employing a single factor. We used the Weighted Least Squares Mean and Variance adjusted (WLSMV) estimator due to the categorical nature of our GULS data (Rosseel, 2012). The robust fit indices were examined to evaluate the model fit (i.e., chi-square test, Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR)).[3] Additionally, we scrutinized the factor loadings (λ) of the items, excluding those with $\lambda \leq .30$ (Tavakol & Wetzel, 2020) and reevaluating the model. Subsequently, McDonald's ω was employed to assess the reliability of the remaining item set after the UVA and CFA analyses, following the commonly used threshold that $\omega \geq .70$ indicates acceptable reliability (Dunn et al., 2014). Moreover, $.60 \leq \omega < .70$ is considered sufficient for decision making at group level, and $.70 \leq \omega < .80$ at individual level (Evers et al., 2009; Nunnally & Bernstein, 1994). Further, we applied Item Response Theory (IRT) to the dichotomous data concerning the retained items following the UVA and CFA analyses. More specifically, we compared the model fits of both

[3] A non-significant chi-square test, CFI above .90, and RMSEA and SRMR values below .06 indicate good fit. RMSEA and SRMR values between .06 and .08 signify acceptable fit (Rosseel, 2012).

the one- and two-parameter IRT models, along with the IRT model incorporating a guessing parameter, using Analysis of Variance tests. The Test Information Curves (TICs) of these IRT models were plotted with a particular focus on clarifying the amount of information provided by the test within the lower range [-6, 0]. At the item level, we calculated the item-total score correlations and examined the item in- and outfit statistics[4], aiming for values between the traditional mean square item fit bounds of 0.75 and 1.33 (Katz et al., 2021). Subsequently, we classified the item difficulty parameters (b) as follows: $b < -2.00$ (very easy), $-2.00 \leq b \leq 2.00$ (moderately difficult), and $b > 2.00$ (very difficult) (Hambleton et al., 1991). The item discrimination parameters (a) were categorized in the following way: $0.01 \leq a \leq 0.34$ (very low), $0.35 \leq a \leq 0.64$ (low), $0.65 \leq a \leq 1.34$ (moderate), $1.35 \leq a \leq 1.69$ (high), $a \geq 1.70$ (very high) (Baker, 2001). Corresponding visualizations were operationalized through Item Characteristic Curves (ICCs).

Second, this data analysis procedure (i.e., UVA, CFA, ω, IRT with TICs, and ICCs) was reiterated on the 2020-2021 to 2022-2023 data (i.e., sample 2), using the recommended set of GULS items derived from the analyses of the preceding three years.

Predictive Validity

First, we examined the predictive validity of GULS for academic achievement *across study programs* on the 2017-2018 to 2019-2020 (i.e., subsample 1) and 2020-2021 to 2022-2023 data (i.e., subsample 2) separately (see also *Participants*). We controlled for sex, SES, mathematics hours secondary education, vocabulary knowledge, and mathematics category. After checking the sample-specific reliability of GULS using McDonald's ω, GULS was included as a total score (i.e., the mean of the retained set of GULS items after the subsample-specific construct validity and reliability analyses). For both subsamples, the descriptives and

[4] Item infit and outfit assess the fit between observed and expected item responses in IRT models, with item infit focusing on responses near the item's difficulty and item outfit considering overall fit across ability levels. Item outfit is also sensitive to outliers.

correlation matrix were examined, followed by a linear regression with academic achievement (i.e., GPA or study success) as the dependent variable and GULS, along with the control variables, as the predictors. The Pearson correlation coefficient was used with the following rule of thumb: $|r|$ = .00 (no correlation), between .00 < $|r|$ < .20 (very weak), .20 ≤ $|r|$ < 0.4 (weak), .40 ≤ $|r|$ < .60 (moderately strong), .60 ≤ $|r|$ < .80 (strong), .80 ≤ $|r|$ < 1.00 (very strong), and $|r|$ = 1.00 (perfect) (Krehbiel, 2004). Further, we focused on GULS's parameter estimates, and its individual and unique explained variance in academic achievement using Adjusted R-squared values.[5] Also, we investigated whether the explained variance in academic achievement differed between the full model and the model without GULS. Multicollinearity was assessed using Variance Inflation Factor (VIF) values for each independent variable. Roughly, VIF values below 10 are acceptable, though values over 5 may indicate substantial multicollinearity, suggesting a threshold of VIF < 5 for reliable regression results (Marcoulides & Raykov, 2019).

Second, we investigated the predictive validity of GULS for academic achievement *per study program* (segmented by math category). We controlled for sex, SES, mathematics hours secondary education, vocabulary knowledge, and mathematical proficiency. To ensure sufficient statistical power per study program, we used the total dataset from 2017-2018 to 2022-2023 (see also *Participants*). Per study program, reliability was verified through McDonald's ω. We computed the descriptives and conducted linear regressions with GPA or study success as the dependent variable, and GULS and the control variables as the independent variables. Again, focus was directed towards GULS's parameter estimates, its individual and unique explained variances (see Footnote 5) in academic achievement using Adjusted R-squared values, and the difference in explained variance in academic achievement

[5] Individual explained variance indicates a predictor's total impact on the dependent variable regardless of other predictors, while unique explained variance shows its distinct contribution (i.e., incremental predictive validity).

between the full models and models without GULS. Multicollinearity was also tested per study program (Marcoulides & Raykov, 2019).

Results

Construct Validity and Reliability

Sample 1

In the dataset spanning from 2017-2018 to 2019-2020, we initially considered 25 GULS items. Upon evaluating their wTO values in relation to other items, the UVA shows the exclusion of GULS2 (wTO = .26 paired with GULS1), GULS21 (wTO = .47 paired with GULS25), and GULS23 (wTO = .38 paired with GULS22). The initial CFA on the remaining 22 GULS items reveals a good model fit (χ^2_{scaled} (209) = 1,728.53, $p < .001$, $\text{CFI}_{\text{robust}} = .901$, $\text{RMSEA}_{\text{robust}} = .049$, SRMR = .039). After removing GULS19 ($\lambda = .24$) and GULS20 ($\lambda = .30$), the second CFA again demonstrates a good model fit (χ^2_{scaled} (170) = 947.69, $p < .001$, $\text{CFI}_{\text{robust}} = .926$, $\text{RMSEA}_{\text{robust}} = .045$, SRMR = .034) but highlights the exclusion of GULS11 ($\lambda = .29$) and GULS18 ($\lambda = .29$). The final CFA confirms the good model fit, with all 18 GULS items showing factor loadings $\lambda > .30$. See Table 2 for the final model fit and reliability measures, and Table 3 for the item factor loadings.

In the IRT analysis, the two-parameter model (AIC = 189,380.00) with 18 GULS items ($M = 14.3$, $SD = 2.8$) outperforms the one-parameter model (AIC = 190,576.30, $p < .001$). The three-parameter model with guessing parameter (AIC = 189,318.10, $p < .001$) demonstrates a better model fit compared with the two-parameter model. However, the two-parameter model offers more test information within the lower range of language proficiency ability [-6, 0] (85% versus 82% attained by the three-parameter model). See Figure 1 for the TIC of the two-parameter model, and Appendix 3A, Figure A1 for the TIC of the three-parameter model. At the item level, all item in-and outfit statistics meet the accepted criteria, falling between 0.75 and 1.33. Item difficulties range from -2.84 to 0.33, and item discriminations vary between 0.62 and 2.11. ICCs are available in Appendix 3A, Figure A2. Plots regarding the item in- and outfits can be found in Figure A3. The item parameters and item-total score correlations are displayed in Table 3.

Sample 2

In the dataset spanning from 2020-2021 to 2022-2023, we considered 18 GULS items. This 18-item set remains consistent following UVA and CFA. No wTO > .25 values are found between two items, the model fit is good with none of the item factor loadings $\lambda \leq .30$. Refer to Table 2 for the model fit and reliability measures, and Table 3 for the item factor loadings.

In the IRT analysis, the two-parameter model (AIC = 274,752.50) using 18 GULS items (M = 13.9, SD = 3.1) demonstrates superior performance compared to the one-parameter model (AIC = 276,750.50, $p < .001$). The three-parameter model (AIC = 274,628.30) incorporating a guessing parameter shows a better model fit compared with the two-parameter model ($p < .001$). However, the two-parameter model provides more test information within the lower range of language proficiency ability [-6, 0] (84% versus 81% obtained by the three-parameter model). Refer to Figure 1 for the TIC of the two-parameter model, and to Appendix 3A, Figure A4 for the TIC of the three-parameter model. At the item level, all in- and outfit statistics align with the target range of 0.75 to 1.33. Item difficulties span from -2.26 to 0.43, while item discriminations range between 0.62 and 2.17. The ICCs can be found in Appendix 3A, Figure A5. Plots regarding the item in- and outfits can be found in Figure A6. The item parameters and their total score correlations are presented in Table 3.

Table 2

Model Fit and Reliability per Sample

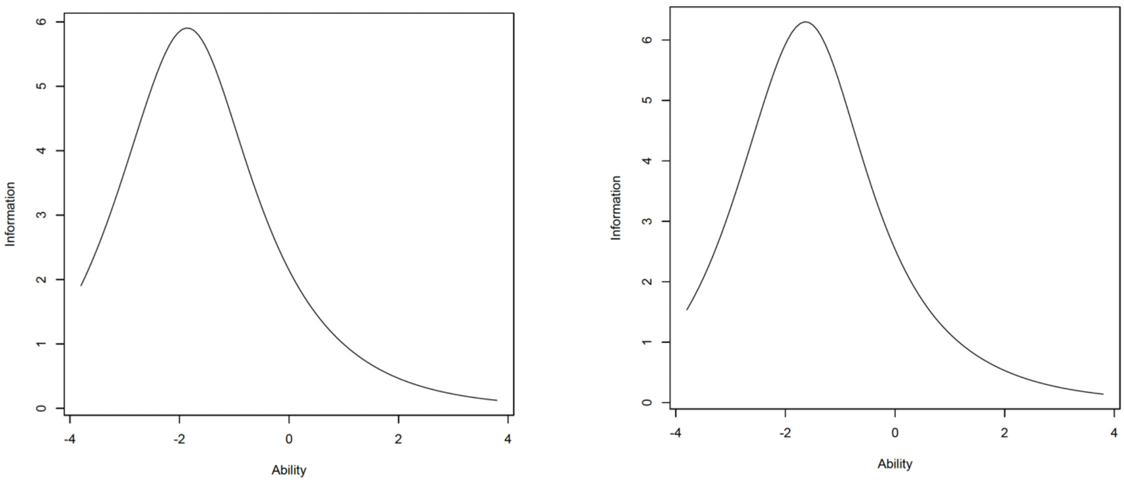
	Sample 1	Sample 2
χ^2_{scaled} (135) (*p*)	684.20 (< .001)	930.44 (< .001)
$\text{CFI}_{\text{robust}}$	.936	.953
$\text{RMSEA}_{\text{robust}}$	.047	.041
SRMR	.032	.029
McDonald's ω	.74	.76

Note. χ^2_{scaled} (df) = chi-square test with degrees of freedom placed in parentheses, $\text{CFI}_{\text{robust}}$ = Robust Comparative Fit Index, $\text{RMSEA}_{\text{robust}}$ = Robust Root Mean Square Error of Approximation, SRMR = Standardized Root Mean Square Residual. The first four rows present the results for the goodness-of-fit measures. The last row pertains to the reliability results. Samples 1 and 2 respectively encompass students from academic years 2017-2018 to 2019-2020 (N_1 = 12,527) and 2020-2021 to 2022-2023 (N_2 = 17,204). The reported values correspond to the 18-item final model.

Figure 1

Test Information Curves Two-Parameter Models per Sample

Sample 1 Sample 2



Note. N_1 = 12,527, N_2 = 17,204. The two-parameter model of sample 1 offers 85% test information within the lower range of language proficiency ability [-6, 0], while the two-parameter model of sample 2 offers 84%.

Predictive Validity

For the predictive validity analyses across study programs and study program-specific, no VIF value exceeds 2, indicating the absence of multicollinearity issues. The predictive validity results of GULS for study success (after the second-chance exam period), across study programs and study-program specific, can be found in Appendix 3B.

Across Study Programs

Subsample 1. For the 2017-2018 to 2019-2020 data across 16 study programs, the correlation matrix can be found in Appendix 3A, Table A2. GULS (M = 14.1, SD = 2.9) has a significant effect on academic achievement (i.e., GPA), when controlled for sex, SES, mathematics hours secondary education, vocabulary knowledge, and mathematics category ($t(8,237) = 18.20$, $p < .001$, $B = 1.20$, $\beta = 0.20$). Individually, the explained variance of GULS in GPA is $R^2 = .06$. GULS's unique contribution to the explained variance in GPA is $R^2 = .04$. The difference in explained variance between the full model ($R^2 = .12$) and the model without GULS is significant ($F_{Change}(1, 8,237) = 331.32$, $p < .001$). The reliability $\omega = .73$.

Subsample 2. For the 2020-2021 to 2022-2023 data across 16 study programs, the correlation matrix is available in Appendix 3A, Table A2. GULS (M = 13.6, SD = 3.1) has a significant effect on GPA, when controlled for sex, SES, mathematics hours secondary education, vocabulary knowledge, and mathematics category ($t(10,884) = 23.87$, $p < .001$, $B = 1.33$, $\beta = 0.22$). Individually, the explained variance of GULS in GPA is $R^2 = .08$. GULS's unique contribution to the explained variance in GPA is $R^2 = .04$. The difference in explained variance between the full model ($R^2 =.17$) and the model without GULS is significant ($F_{Change}(1, 10,884) = 569.85$, $p < .001$). The reliability $\omega = .75$.

Study Program-Specific

The data from 2017-2018 to 2022-2023 show that GULS significantly predicts GPA for all the basic mathematics study programs (M range [13.2, 15.2], SD range [2.2, 3.1]), when controlled for sex, SES, mathematics hours secondary education, vocabulary knowledge and mathematical proficiency (i.e., for each study program, $p < .001$). Individually, the explained variance of GULS in GPA varies from $R^2 = .07$ (Physical Therapy and Motor Rehabilitation, and Political Sciences) to $R^2 =.21$ (Applied Language Studies). GULS's unique contribution to the explained variance in GPA ranges from $R^2 = .02$ (Physical Therapy and Motor Rehabilitation) to $R^2 = .11$ (Applied Language Studies). The differences in explained variance between the full models and the models without GULS are significant (p's $< .001$).

Across the advanced mathematics study programs (*M* range [12.7, 14.9], *SD* range [2.7, 3.3]), the study programs Biochemistry and Biotechnology, and Engineering Technology diverge from the observed pattern where GULS significantly predicts GPA, when controlled for sex, SES, mathematics hours secondary education, vocabulary knowledge and mathematical proficiency (*p* range [< .001, .260]). Individually, the explained variance of GULS in GPA varies from $R^2 = .02$ (Engineering Technology) to $R^2 = .08$ (Business Administration and (Applied/Business) Economics). GULS's unique contribution to the explained variance in GPA ranges from $R^2 < .01$ (Biochemistry and Biotechnology, and Engineering Technology) to $R^2 = .05$ (Business Administration). The differences in explained variance between the full models and the models without GULS are significant, except for the study programs 'Biochemistry and Biotechnology', and 'Engineering Technology' (*p* range [< .001, .260]). For details on GULS's descriptives, predictive validity measures for GPA, reliability, and the total explained variance of the full model per study program, we refer to Table 4.

Table 3

Item Factor Loadings, Difficulty and Discrimination Parameters and Item-Total Correlations per Sample

Item	Factor Loading		Item Difficulty		Item Discrimination		Item-Total Correlation	
	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2
GULS1	.59	.57	-2.28	-2.06	1.33	1.22	.40	.42
GULS3	.52	.51	-0.11	0.20	1.01	0.99	.49	.47
GULS4	.36	.37	0.33	0.43	0.62	0.64	.39	.39
GULS5	.40	.42	-1.60	-1.37	0.75	0.79	.40	.41
GULS6	.54	.50	-1.53	-1.52	1.11	1.01	.46	.44
GULS7	.65	.68	-2.20	-2.05	1.53	1.59	.42	.45
GULS8	.35	.33	0.01	0.28	0.65	0.62	.39	.37
GULS9	.59	.60	-1.95	-1.87	1.27	1.30	.44	.45
GULS10	.41	.43	-1.96	-1.96	0.76	0.79	.38	.38
GULS12	.55	.55	-1.95	-1.79	1.14	1.14	.43	.44
GULS13	.57	.56	-1.58	-1.52	1.19	1.16	.46	.46
GULS14	.52	.55	-1.57	-1.42	1.08	1.15	.45	.47
GULS15	.49	.58	-2.84	-2.26	1.04	1.23	.33	.40
GULS16	.59	.66	-2.49	-1.91	1.35	1.53	.37	.46
GULS17	.50	.50	-1.84	-1.80	1.00	0.97	.42	.42
GULS22	.67	.68	-1.45	-1.14	1.52	1.59	.51	.54
GULS24	.77	.78	-1.79	-1.52	2.11	2.17	.50	.55
GULS25	.51	.56	-1.92	-1.59	1.03	1.17	.41	.45

Note. Samples 1 and 2 respectively encompass students from academic years 2017-2018 to 2019-2020 (N_1 = 12,527) and 2020-2021 to 2022-2023 (N_2 = 17,204). Exclusions after UVA comprised GULS2, GULS21, and GULS22. Following CFA's, exclusions involved GULS11, GULS18, GULS19, and GULS20. Further details are available in *Method*.

Table 4

GULS's Descriptives, Predictive Validity Measures for GPA, Reliability, and Total Explained Variance Full Model per Study Program

	M (*SD*)	*t*	*B* (*b*)	R^2			ω
				Individual	Unique	Total	
Basic mathematics							
Applied Language Studies	14.7 (2.6)	9.11***	2.79 (0.31)	.21	.11***	.28	.70
Communication Sciences	14.2 (2.6)	4.63***	1.38 (0.30)	.09	.04***	.13	.68
Criminological Sciences	13.5 (2.8)	7.50***	1.37 (0.18)	.09	.04***	.16	.68
Educational Sciences	13.9 (2.6)	7.69***	1.38 (0.18)	.11	.08***	.19	.63
History	15.2 (2.2)	4.12***	1.94 (0.47)	.12	.03***	.18	.64
Law	14.3 (2.7)	13.28***	1.87 (0.14)	.15	.07***	.22	.72
Pharmaceutical Sciences	13.7 (3.1)	6.58***	1.19 (0.18)	.11	.03***	.19	.77
Physical Therapy and Motor Rehabilitation	13.2 (3.1)	6.46***	0.80 (0.12)	.07	.02***	.20	.74
Political Sciences	14.5 (2.5)	4.55***	1.57 (0.35)	.07	.04***	.10	.68
Psychology	13.8 (2.8)	12.73***	1.55 (0.12)	.11	.05***	.22	.69
Advanced mathematics							
Biochemistry and Biotechnology	14.1 (3.1)	1.18	0.34 (0.29)	.03	<.01	.14	.78
Biomedical Sciences	13.7 (3.1)	4.10***	0.82 (0.20)	.06	.01***	.14	.76
Bioscience Engineering	14.9 (2.7)	3.29**	0.60 (0.18)	.07	.01**	.24	.78
Business Administration	12.7 (3.3)	10.43***	1.25 (0.12)	.08	.05***	.15	.75
(Applied/Business) Economics	13.9 (3.1)	7.70***	0.98 (0.13)	.08	.03***	.17	.78
Engineering Technology	13.4 (3.2)	1.13	0.17 (0.15)	.02	<.01	.16	.77

Note. **p* < .05, ***p* < .01, ****p* < .001. ω = McDonald's ω as reliability measure. R^2 = Adjusted R-squared. Total R^2 represents the explained variance by the full model, whereas unique R^2 quantifies the specific contribution of GULS to GPA, distinct from other predictors (i.e., sex, SES, mathematics hours secondary education, vocabulary knowledge, and mathematical proficiency). Meanwhile, individual R^2 reflects the overall contribution of GULS to GPA, regardless of additional predictors. The advanced mathematics study programs are associated with more advanced mathematical curricula. Through the SIMON project (Fonteyne, 2017), students enrolling in basic mathematics study programs complete a less difficult math assessment compared to students enrolling in more advanced mathematics study programs.

Discussion

Academic language proficiency (hereafter: language proficiency) is particularly important for students entering higher education (Knoch & Elder, 2013; Read, 2016). Indeed, language proficiency, assessed for admission eligibility (e.g., Clinton-Lisell et al., 2022; Ihlenfeldt & Rios, 2023) or through post-entry language assessment (PELA) (e.g., Heeren et al., 2021; van Dijk, 2015), shows predictive validity for academic achievement (Elder, 2017; Knoch & Elder, 2013).

The use of PELAs is relevant for higher education institutions with low or no admission requirements, as these assessments help identify at-risk students and provide them with the necessary language support to improve their academic trajectory (Elder, 2017; Knoch & Elder, 2013; Read, 2016). However, few PELAs are developed for languages other than English, like Dutch. An existing Dutch PELA, constructed by Heeren and colleagues (2021) certainly has merits, but the test and its detailed construct validation evidence are not publicly accessible to the best of our knowledge. The researchers also used a less conventional, however still valuable, metric of academic achievement compared to Grade Point Average (GPA) to examine the predictive validity of language proficiency for academic achievement. Moreover, the differential contribution of language proficiency to academic achievement for various academic disciplines is acknowledged (Elder, 2017; Read, 2016), yet often not implemented.

The present study aimed to enhance PELA in higher education by introducing and validating the Ghent University Language Screening (GULS) as a fully open access (i.e., easy-to-administer, free and publicly available) PELA. Concretely, GULS assesses Dutch language proficiency, in terms of reading comprehension, of first-year higher education students. First, GULS's construct validity and reliability were evaluated and confirmed through data from the academic years 2017-2018 to 2019-2020 (N_1 = 12,527) and 2020-2021 to 2022-2023 (N_2 = 17,204). Second, GULS's predictive validity for academic achievement (i.e., GPA and study success) was assessed on data over these two three-year periods across 16 study programs (n_1 = 8,244; n_2 = 10,891), succeeded by predictive validity analyses for each study program across the combined six-year span.

Construct Validity and Reliability

The results concerning construct validity indicate that GULS with 18 items is the optimal version to avoid item redundancy (Christensen et al., 2023) and ensure sufficiently high item factor loadings (Tavakol & Wetzel, 2020). Data from the periods 2017-2018 to 2019-2020 and 2020-2021 to 2022-2023 subsequently fit well with the one-factor model comprising 18 items, where the factor represents language proficiency in terms of reading comprehension (Kintsch, 2013; OECD, 2023). Additionally, GULS appears to provide the most test information at the lower end of the language proficiency spectrum. In other words, GULS is particularly sensitive and insightful for students with lower language proficiency scores, enabling the identification and assistance of those students in need of additional language development support. This finding corresponds with the objectives of PELAs (Elder, 2017; Knoch & Elder, 2013; Read, 2016). Especially in low-stakes higher education environments with a more diverse enrolling student population, the ability to detect and support such at-risk students at an early stage is key to be able to achieve the so-called equity for equal achievement (Espinoza, 2007).

GULS's reliabilities across study programs exceed the rule of thumb. Hence, GULS serves multiple purposes effectively (Evers et al., 2009; Nunnally & Bernstein, 1994). First, the test can be used in population-level research across and within higher education study programs, including both basic as well as advanced mathematics study programs. Second, non-binding advice can be offered to (prospective) first-year higher education students regarding their language proficiency in terms of reading comprehension. Third, tailoring this individual advice to specific advanced mathematics study programs is possible based on GULS's reliabilities. However, caution is warranted for basic mathematics study programs, as the empirical evidence regarding the reliabilities is inconclusive as such in six out of the 10 included study programs (although, they do satisfy the threshold for population-level research) (Evers et al., 2009; Nunnally & Bernstein, 1994).

Moreover, the construct validity and reliability findings across study programs are confirmed by a second, independent prospective dataset from 2021-2022 to 2022-2023. This fact reinforces the robustness of GULS as a valid and reliable tool for assessing Dutch language proficiency, in terms of reading comprehension, of first-year students in higher education.

Predictive Validity

Consistent across the datasets from 2017-2018 to 2019-2020 and from 2020-2021 to 2022-2023, GULS is a significant predictor of academic achievement (i.e., GPA and study success) across 16 study programs. Language proficiency measured through GULS contributes modestly to the prediction of academic achievement, aligning with the correlations found in previous research using language assessment for admission in English-medium higher education institutions (for meta-analyses, see e.g., Ihlenfeldt & Rios, 2023; Clinton-Lisell et al., 2022) or (non-)Dutch PELAs (e.g., De Wachter et al., 2013; Heeren et al., 2021; van Dijk, 2015). Researchers also emphasize that language proficiency typically explains no more than 10% of the variance in academic achievement (Elder, 2017; Knoch & Elder, 2013). We further acknowledge that language proficiency is a relevant yet not exclusive determinant of academic achievement. The interplay between (non-)cognitive and personal/educational background characteristics is indeed essential for accurately predicting academic achievement (Elder, 2017; Fonteyne et al., 2017; Richardson et al., 2012; Read, 2016). Importantly, even when controlling for personal and educational background characteristics, vocabulary knowledge, and mathematics category in the present study, we observe incremental predictive validity of GULS for academic achievement. This finding implies that language proficiency assessed through GULS continues to capture unique aspects of and provide insights into academic achievement that are not entirely covered by other well-established predictors for academic achievement like SES. Literature indicates that personal background characteristics rather weakly contribute to academic achievement beyond language proficiency and do not impact the incremental predictive validity of language proficiency, unlike educational background characteristics (Heeren et al., 2021). In the present study, accounting for mathematics hours in secondary education as an educational background characteristic also increases the total variance explained in GPA across study programs. At the same time, however, this educational background characteristic does not alter the unique contribution of language proficiency as measured by GULS to academic achievement.[6] Researchers' observation that educational background characteristics are primarily reflected in language proficiency (Heeren et al., 2021; Stricker, 2004) does not emerge prominently in our context across study programs.

In addition, study-program specific analyses spanning the six academic years likewise indicate that GULS modestly predicts academic achievement across all basic mathematics study programs. Similarly to the overall analysis, after accounting for personal and educational background characteristics, vocabulary knowledge, and mathematical proficiency, GULS also shows incremental predictive validity for academic achievement. The models including GULS are regarded superior to the models without GULS for predicting academic achievement. Notably, the role of language proficiency in academic achievement prediction

---

[6] The difference in unique explained variance of GULS with or without mathematics hours secondary education in the regression model (including the other control variables) amounts 1% for both GPA as well as study success, on average across the two three-year periods.

may differ by discipline (Elder, 2017; Knoch & Elder, 2013; Read, 2016), regardless of the comparable level of mathematics proficiency that study programs expect from enrolling students. Indeed, the present study confirms a differential contribution of GULS to academic achievement. For example, the Applied Language Studies study program stands out in terms of GULS's predictive validity for academic achievement, with an individual contribution of 21% and unique contribution of 11% in GPA (18% and 10% in study success). Given the study program's focus on language, culture and its practical applications to real-world contexts (Berns & Matsuda, 2006), this outcome is not surprising and illustrates GULS's construct validity well. By contrast, in the discipline of Physical Therapy and Motor Rehabilitation within the basic mathematics study programs, GULS only accounts for 2% of the variance in GPA and 1% in study success when controlled for the other predictors. Yet, the inclusion of GULS still renders the models more effective for predicting academic achievement than the models without GULS. Furthermore, the educational background variable of mathematics hours in secondary education contributes to 7% of the unique variance in both GPA as well as study success within this specific study program. Nevertheless, the impact of this educational background variable on the unique explained variance of GULS is minimal, as this educational background variable only reduces the unique explained variance of GULS by 1% in both GPA and as well as study success when included in the model.

The same pattern regarding the modest (incremental) predictive validity of GULS for academic achievement applies to most advanced mathematics study programs as well. The control variable, mathematics category, indeed shows no effect on academic achievement when controlling for other predictors in the models across study programs. Overall, even for students who choose a study program that is characterized by more advanced mathematical curricula, language proficiency as measured through GULS remains important for their academic achievement. In fact, language and mathematical proficiency are established predictors of academic achievement (e.g., Fonteyne et al., 2017). Moreover, meta-analytic evidence shows a mutual moderate relationship between these proficiencies, primarily attributed to domain-general processes such as executive functions (Ünal et al., 2023). In addition, GULS proves to be sensitive to the role of language proficiency in predicting academic achievement, as its predictive validity for academic achievement also varies by discipline within the advanced mathematics study program group (Elder, 2017; Knoch & Elder, 2013; Read, 2016). However, for the advanced mathematics study programs 'Biochemistry and Biotechnology', and 'Engineering Technology', we find no effect of GULS on academic achievement. This result may be due to these study programs' reliance on more deeply entrenched technical and specific academic language, while GULS is developed to evaluate more general academic language proficiency. Looking further, we also observe that the mathematical proficiency in the study program 'Engineering Technology' and the mathematics hours in secondary education in 'Biochemistry and Biotechnology' explain the most in their prediction models for academic achievement, respectively accounting for 5% and 3% of the variance in GPA and for 5% and 2% of the variance in study success.

Implications

To the best of our knowledge, the present study addresses some gaps in current literature. First, the existence of GULS fills a void concerning the availability of Dutch PELAs, while also being fully open access (i.e., easy-to-administer, free, and publicly available). Second, we reported in detail on the construct validity and reliability, thereby enhancing transparency, replicability, and reproducibility (Min & Aryadoust, 2021), and facilitating public access to the psychometric properties of a PELA (Knoch & Elder, 2013). Third, we examined GULS's (incremental) predictive validity for academic achievement both across as well as within various study programs (Elder, 2017; Hauspie et al., 2024; Read, 2016). As a result, the present study provides insights into the overall and differential contribution(s) of language proficiency to academic achievement, including linguistic, social science, economic, and STEM-oriented disciplines, among others.

The present study validated GULS as a Dutch PELA that offers practical benefits due to its open access, enabling its ease of deployment in research and practice contexts. With demonstrated construct validity and reliability, GULS proves useful for conducting population-level research across and within diverse study programs (Evers et al., 2009; Nunnally & Bernstein, 1994). For instance, GULS can be utilized to control for Dutch language proficiency when examining academic achievement in future educational research. Moreover, GULS facilitates the provision of individualized, non-binding advice to (prospective) first-year higher education students regarding their Dutch language proficiency levels. This personalized guidance can help identify at-risk students and offer targeted language support interventions. Notably, GULS also allows for the specification of advice for advanced mathematics study programs, although caution is recommended for basic mathematical programs (Evers et al., 2009; Nunnally & Bernstein, 1994). In any case, it is always encouraged to verify reliability within one's own sample in new research (Graham, 2015; Harris, 2003).

Strengths, Limitations and Future Research

GULS assesses language proficiency in terms of reading comprehension, beyond simple vocabulary knowledge, to identify at-risk students, offer language support and enrich their academic journey. We opted to focus on reading comprehension, given its paramount importance during the early stages of higher education (De Wachter et al., 2013; Jansen et al., 2022; Van Houtven et al., 2010). Consequently, GULS does not involve other language proficiency skills such as listening, which can be interesting to address in future research and application.

For the construct/predictive validity and reliability analyses across study programs, large sample sizes were used, with initial results from a three-year period subsequently confirmed by data from an additional three years across study programs. To ensure robust statistical power for the program-specific predictive validity analyses, we utilized the student data over the six years. On this matter, to examine the predictive validity of GULS for academic achievement, we considered both GPA and study success as academic achievement measures. As such, we address operationalizations of academic achievement that are applicable to both European (i.e., study success) and American educational contexts (i.e., GPA). Future (longitudinal) research could also investigate GULS's predictive validity for timely bachelor's degree completion, allowing for a more comprehensive assessment of how language proficiency impacts students' academic progress and degree attainment over time.

In addition, we accounted for a variety of control variables in the predictive validity analyses, including personal and educational background characteristics and cognitive factors. Non-cognitive factors, however, were not considered in the present study, highlighting a direction for future research as much of the variance in academic achievement remains unexplained. Moreover, attention could also be given to (theory-based) interactions between language proficiency and various other variables. Concerning the educational background characteristics furthermore, we did not incorporate the more typically used last-year high school GPA. In fact, last-year high school GPA is challenging to compare across schools, especially in the present study's context in which no standardized exams are held in the last year of secondary education. Also, this variable would have required self-reporting by students due to ethical and privacy concerns, potentially leading to bias from social desirability. The number of mathematics hours in secondary education was therefore regarded as a more objective indicator of educational background, although students' answers also depended on self-reporting.

Conclusion

In the present study we validated GULS, a Dutch post-entry language assessment with fully open access. More specifically, GULS assesses reading comprehension of first-year students in higher education. GULS is proven to be valid and reliable, particularly for identifying students who require language support at the start of higher education, and predicts academic achievement. The accessibility of GULS enhances its utility as a Dutch language proficiency test for future population-level research and for providing advice to (prospective) higher education students.

References

Abunawas, M. (2014). *A meta-analytic investigation of the predictive validity of the Test of English as a Foreign Language (TOEFL) scores on GPA*. https://oaktrust.library.tamu.edu/bitstream/handle/1969.1/154156/ABUNAWAS-DISSERTATION-2014.pdf?sequence=1

Baker, F. B. (2001). *The Basics of Item Response Theory. Second Edition.* MD: ERIC Clearinghouse on Assessment and Evaluation.

Berns, M., & Matsuda, K. (2006). Applied Linguistics: Overview and History. In *Encyclopedia of Language & Linguistics* (pp. 394–405). Elsevier. https://doi.org/10.1016/B0-08-044854-2/00599-X

Christensen, A. P., Garrido, L. E., & Golino, H. (2023). Unique Variable Analysis: A Network Psychometrics Method to Detect Local Dependence. *Multivariate Behavioral Research*, *58*(6), 1165–1182. https://doi.org/10.1080/00273171.2023.2194606

Clinton-Lisell, V., Taylor, T., Carlson, S. E., Davison, M. L., & Seipel, B. (2022). Performance on Reading Comprehension Assessments and College Achievement: A Meta-Analysis. *Journal of College Reading and Learning*, *52*(3), 191–211. https://doi.org/10.1080/10790195.2022.2062626

De Wachter, L., Heeren, J., Marx, S., & Huyghe, S. (2013). "Taal: een noodzakelijke, maar niet de enige voorwaarde tot studiesucces: De correlatie tussen de resultaten van een taalvaardigheidstoets en de slaagcijfers bij eerstejaarsstudenten aan de KU Leuven." [Language Proficiency: A Necessary, but Not the Only, Condition for Study Success: A Correlation between the Results of a Language Proficiency Test and Academic Achievement of First-Year Students.]. *Levende Talen Tijdschrift*, *14*(4), 28–36. https://lt-tijdschriften.nl/ojs/index.php/ltt/article/view/549

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399–412. https://doi.org/10.1111/bjop.12046

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (2024). *Ethnologue: Languages of the World. Twenty-seventh edition.* Texas: SIL International. https://www.ethnologue.com/language/nld/

Elder, C. (2017). Language Assessment in Higher Education. In *Language Testing and Assessment* (pp. 271–286). Springer International Publishing. https://doi.org/10.1007/978-3-319-02261-1_35

Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: a new model for analysis of the educational process. *Educational Research*, *49*(4), 343–363. https://doi.org/10.1080/00131880701717198

European Commission, D.-G. for E. Y. S. and C. (2015). *ECTS users' guide 2015*. https://data.europa.eu/doi/10.2766/87192

Evers, A., Lucassen, W., Meijer, R., & Sijtsma, A. (2009). *COTAN beoordelingssysteem voor de kwaliteit van tests (geheel herziene versie)*. http://www.psynip.nl/website/wat-doethet- nip/tests/beoordelingsprocedure/beoordelingsprocedure

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.

Fonteyne, L., De Fruyt, F., Dewulf, N., Duyck, W., Erauw, K., Goeminne, K., Lammertyn, J., Marchant, T., Moerkerke, B., Oosterlinck, T., & Rosseel, Y. (2015). Basic mathematics test predicts statistics achievement and overall first year academic success. *European Journal of Psychology of Education*, *30*(1), 95–118. https://doi.org/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. https://doi.org/10.1016/j.lindif.2017.05.003

Gagen, T. (2019). *The predictive validity of IELTS scores: A meta-analysis*. https://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=8762&context=etd

Graham, S. (2015). Inaugural editorial for the Journal of Educational Psychology. *Journal of Educational Psychology*, *107*(1), 1–2. https://doi.org/10.1037/edu0000007

Hambleton, R., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. CA: Sage Publications.

Harris, K. R. (2003). Editorial: Is the work as good as it could be? *Journal of Educational Psychology*, *95*(3), 451–452. https://doi.org/10.1037/0022-0663.95.3.451

Hauspie, C., Schelfhout, S., Dirix, N., Fonteyne, L., Janse, M., Szmalec, A., Vereeck, A., & Duyck, W. (2024). Does Studying Latin in Secondary Education Predict Study Achievement in Academic Higher Education? *Language Learning*. https://doi.org/10.1111/lang.12639

Heeren, J., Speelman, D., & De Wachter, L. (2021). A practical academic reading and vocabulary screening test as a predictor of achievement in first-year university students:

implications for test purpose and use. *International Journal of Bilingual Education and Bilingualism*, *24*(10), 1458–1473. https://doi.org/10.1080/13670050.2019.1709411

Hulstijn, J. H. (2015). *Language Proficiency in Native and Non-native Speakers*. John Benjamins B.V.

Ihlenfeldt, S. D., & Rios, J. A. (2023). A meta-analysis on the predictive validity of English language proficiency assessments for college admissions. *Language Testing*, *40*(2), 276–299. https://doi.org/10.1177/02655322221112364

Jansen, C., De Wachter, L., Van Dun, P., & Frik, T. (2022). *Taalcompetentie in het Nederlands van Nederlandstalige studenten in het hoger onderwijs in Nederland en Vlaanderen*. https://taalunie.org/publicaties/213/taalcompetentie-in-het-nederlands-van-nederlandstalige-studenten-in-het-hoger-onderwijs-in-nederland-en-vlaanderen

Katz, D., Clairmont, A., & Wilton, M. (2021). *Measuring what Matters: Introduction to Rasch Analysis in R*. https://bookdown.org/dkatz/Rasch_Biome/

Kennedy, T. J., & Sundberg, C. W. (2020). *21st Century Skills* (pp. 479–496). https://doi.org/10.1007/978-3-030-43620-9_32

Kintsch, W. (2013). Revisiting the Construction–Integration Model of Text Comprehension and Its Implications for Instruction. In *Theoretical Models and Processes of Reading* (pp. 807–839). International Reading Association. https://doi.org/10.1598/0710.32

Knoch, U., & Elder, C. (2013). A framework for validating post-entry language assessments (PELAs). *Papers in Language Testing and Assessment*, *2*(2), 48–66. https://arts.unimelb.edu.au/__data/assets/pdf_file/0007/1771306/4_Knoch_Elder_0.pdf

Krehbiel, T. C. (2004). Correlation Coefficient Rule of Thumb. *Decision Sciences Journal of Innovative Education*, *2*(1), 97–100. https://doi.org/10.1111/j.0011-7315.2004.00025.x

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Marcoulides, K. M., & Raykov, T. (2019). Evaluation of Variance Inflation Factors in Regression Models Using Latent Variable Modeling Methods. *Educational and Psychological Measurement*, *79*(5), 874–882. https://doi.org/10.1177/0013164418817803

Min, S., & Aryadoust, V. (2021). A systematic review of item response theory in language assessment: Implications for the dimensionality of language ability. *Studies in Educational Evaluation*, *68*, 100963. https://doi.org/10.1016/j.stueduc.2020.100963

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory (3rd ed.)*. New York: McGraw-Hill.

OECD. (2019). *An OECD Learning Framework 2030* (pp. 23–35). https://doi.org/10.1007/978-3-030-26068-2_3

OECD. (2023). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. https://doi.org/https://doi.org/10.1787/53f23881-en

Read, J. (2015). *Assessing English Proficiency for University Study*. Palgrave Macmillan UK. https://doi.org/10.1057/9781137315694

Read, J. (2016). *Some Key Issues in Post-Admission Language Assessment* (pp. 3–20). https://doi.org/10.1007/978-3-319-39192-2_1

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. https://doi.org/10.1037/a0026838

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2). https://doi.org/10.18637/jss.v048.i02

Schelfhout, S. (2019). *The Influence of Study Interests and (Non-)Cognitive Predictors on Study Choice and Study Achievement in Flemish Higher Education* [Dissertation]. Ghent University.

Siegelman, N., Schroeder, S., Acartürk, C., Ahn, H.-D., Alexeeva, S., Amenta, S., Bertram, R., Bonandrini, R., Brysbaert, M., Chernova, D., Da Fonseca, S. M., Dirix, N., Duyck, W., Fella, A., Frost, R., Gattei, C. A., Kalaitzi, A., Kwon, N., Lõo, K., … Kuperman, V. (2022). Expanding horizons of cross-linguistic research on reading: The Multilingual Eye-movement Corpus (MECO). *Behavior Research Methods*, *54*(6), 2843–2863. https://doi.org/10.3758/s13428-021-01772-6

Stricker, L. J. (2004). The performance of native speakers of English and ESL speakers on the computer-based TOEFL and GRE General Test. *Language Testing*, *21*(2), 146–173. https://doi.org/10.1191/0265532204lt279oa

Tavakol, M., & Wetzel, A. (2020). Factor Analysis: a means for theory and instrument development in support of construct validity. *International Journal of Medical Education*, *11*, 245–247. https://doi.org/10.5116/ijme.5f96.0f4a

Ünal, Z. E., Greene, N. R., Lin, X., & Geary, D. C. (2023). What Is the Source of the Correlation Between Reading and Mathematics Achievement? Two Meta-analytic Studies. *Educational Psychology Review*, *35*(1), 4. https://doi.org/10.1007/s10648-023-09717-5

van Dijk, T. (2015). Tried and tested. *Tijdschrift Voor Taalbeheersing*, *37*(2), 159–186. https://doi.org/10.5117/TVT2015.2.VAND

Wingate, U. (2015). *Academic Literacy and Student Diversity*. Multilingual Matters. https://doi.org/10.21832/9781783093496

Wongtrirat, R. (2010). *English language proficiency and academic achievement of international students: A meta-analysis*. https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1183&context=efl_etds

York, T. Y., Gibson, C., & Rankin, S. (2015). Defining and Measuring Academic Success. *Practical Assessment, Research & Evaluation*, *20*(5), 1–20. http://pareonline.net/getvn.asp?v=20&n=5

4

The Role of Feedback Self-Efficacy in Student Feedback Engagement

Submitted to Assessment & Evaluation in Higher Education, 17th of June 2024.

Accepted for Review, 17th of June 2024.

Mona Bassleer, Stijn Schelfhout, Lot Fonteyne, Wouter Duyck & Nicolas Dirix.

Abstract

Recently, educational feedback research is shifting away from examining effective feedback provision towards focusing more on feedback reception. However, theory-based empirical studies on determinants and underlying mechanisms of student feedback engagement are still needed to develop more targeted interventions. In higher education, addressing student feedback engagement can streamline students' study trajectory and ultimately reduce fail and dropout rates. The present study concentrated on first-year university students receiving feedback on their program-specific validated academic achievement prediction, based on various background/(non-)cognitive variables. This feedback also recommended student-specific remediation and competence training activities. Surveying student feedback engagement yielded longitudinal data of $N = 392$. We analyzed students' intentional and behavioral feedback engagement, their feedback self-efficacy and their received feedback, based on a Theory of Planned Behavior model. The results show that intentional feedback engagement positively influences behavioral feedback engagement, and feedback self-efficacy positively affects intentional feedback engagement. Also, feedback indicating a (fairly) high versus a (very) low chance of study success increases feedback self-efficacy. Furthermore, intentional feedback engagement fully mediates the relationship between feedback self-efficacy and behavioral feedback engagement, as does feedback self-efficacy in the relationship between students' received feedback and intentional feedback engagement. We discuss the value of directing educational interventions towards enhancing feedback self-efficacy as a means of promoting student feedback engagement.

Introduction

Making a higher education study choice that is feasible and aligns with one's vocational interests is challenging for students, especially in an open access higher education system with a wide range of study choice. Indeed, among the full-time students enrolling in a bachelor's program, 12% to 16% drop out after one year (OECD, 2022; Statistiek Vlaanderen, 2024) and first-year fail rates can increase to 60% (Schelfhout et al., 2022), entailing costs for both students and society (OECD, 2022). Therefore, society and students are likely to benefit from adequate support from the beginning of students' higher education journey. This support includes feedback on estimated (first-year) academic success, with recommendations to actions that students can undertake to improve their higher education chances. Noteworthy, such feedback surpasses the typical research focus on feedback related to a specific competence, as elucidated by meta-analyses (Kluger & DeNisi, 1996; Wisniewski et al., 2020).

Feedback has become a cornerstone of educational practice and policy (Hattie & Timperley, 2007; Panadero, 2023). Well-established meta-analyses state the importance of seeing feedback as a multifaced concept, encompassing various forms with distinct effects on diverse student outcomes (Kluger & DeNisi, 1996; Wisniewski et al., 2020). Until lately, feedback effectiveness research predominantly concentrated on identifying key attributes of feedback design (e.g., regarding the feedback type and source) (Wisniewski et al., 2020). As such, however, feedback *provision* is underscored while feedback *reception* is overshadowed, concerning the approach and the extent to which students engage with received feedback (Boud & Molloy, 2013; Van der Kleij & Lipnevich, 2020). Moreover, previous studies consistently demonstrate student dissatisfaction with received feedback, which stresses the relevance to investigate not just the provision but also the reception of feedback (e.g., Mulliner & Tucker, 2017).

Furthering this dialogue, recent efforts rightly highlight the socio-constructivist approach towards feedback where the learner's co-responsibility in the feedback process is emphasized (Dann, 2017; Winstone et al., 2019). Student feedback engagement plays a vital role in closing the so-called feedback gap (Adams et al., 2020), as student feedback engagement is considered paramount to ensure the effectiveness of feedback for student outcomes (Ajjawi & Boud, 2017; Van der Kleij & Lipnevich, 2020). Meanwhile, advancements to describe feedback processes and contributing variables to student feedback engagement are realized (e.g., Lipnevich & Panadero, 2021). However, more empirical understanding of student feedback engagement, and its determinants and underlying mechanisms (i.e., related to how and why feedback works), is still required to elevate our grasp and refine educational practices within this realm (Panadero & Lipnevch, 2022; Winstone & Nash, 2023). Additionally, focusing particularly on student characteristics as possible determinants is advised, thereby recognizing the feedback agent as central to the feedback process (Panadero, 2023). More specifically, students' higher education journey can be enhanced by pinpointing factors that determine student feedback engagement and unraveling its mechanisms, allowing for a more targeted approach. Especially those students at risk of low predicted study success should be reached and encouraged to engage with the feedback they receive. Ultimately, (first-year) fail and drop-out rates can be reduced, benefiting both students and society (OECD, 2022; Schelfhout et al., 2022).

The present study addresses the need for more theory-based empirical research on determinants and underlying mechanisms related to student feedback engagement (Panadero, 2023; Winstone & Nash, 2023). For this purpose, we consider the Theory of Planned Behavior (TPB) (Ajzen, 1991, 2012), combined with insights from prior feedback research, to examine student feedback engagement within the higher education context. The TPB predicts and explains diverse behaviors by emphasizing the role of behavioral intention, influenced by individual characteristics like perceived behavioral control (for meta-analyses, see e.g., Hirschey et al., 2020; Riebl et al., 2015). In contrast to existing feedback models (Lipnevich & Panadero, 2021), the TPB is in fact broadly recognized as a consistent theoretical model. Hence, the TPB has become widely spread in fields like business and public environmental health (Bosnjak et al., 2020), and more recently also within educational contexts (e.g., Opoku et al., 2021). The present study thus adopts a TPB-based model in which the behavior to be predicted and explained corresponds to student feedback engagement within the higher education context (see also Figure 1). This approach aligns with socio-constructivist feedback principles (Dann, 2017; Winstone et al., 2019), and offers insights into student characteristics as determinants of student feedback engagement and its underlying mechanisms that can be targeted in educational interventions aimed at promoting student feedback engagement (Ajjawi & Boud, 2017; Winstone et al., 2019). Consequently, students (at-risk of failure) are enabled to receive more tailored support during their higher education journey, which can eventually facilitate their study careers.

Feedback

Feedback literature encompasses various viewpoints and interpretations of the term feedback (Winstone & Boud, 2022). We adopt a feedback definition that aligns to that of Lipnevich and Panadero (2021), which is compiled from prominent feedback models and theories like Hattie and Timperley's (2007) feedback model. Feedback is information on students' present performance level, goal alignment, future direction, and improvement guidance. The information originates from various sources, such as teachers, students or, as in the present study, computer-based systems (see also Footnote 2) (Fonteyne, 2017). Additionally, feedback aims to improve student outcomes by encouraging students' active processing of the performance-related information (Lipnevich & Panadero, 2021). Within the TPB, feedback can be considered an information-related background characteristic influencing intention and behavior antecedents (see further) (Ajzen, 2020; Ajzen & Fishbein, 2005).

The present study's feedback emanates from a self-assessment tool that probes academic potential and vocational interests, aiding prospective students in their decision-making process of their study choice, and offering tailored guidance to first-year higher education students (i.e., the SIMON project). Instead of manipulating feedback conditions experimentally and/or virtually (Wisniewski et al., 2020), students are presented with their actual first-year predicted chance of study success (i.e., (very) low, (fairly) high). These predictions are validated based on historical longitudinal data of background and (non-)cognitive predictors of academic achievement. Additionally, students receive a comprehensive overview of their (non-)cognitive competences, linked to recommendations for remediation/competence training activities (Fonteyne, 2017). As such, this feedback extends beyond feedback focused on a specific competence, as commonly observed in feedback research (Kluger & DeNisi, 1996; Wisniewski et al., 2020).

Student Feedback Engagement

Student feedback engagement has reached consensus in literature as a three-dimensional construct, as the concept features a cognitive, affective, and behavioral dimension (e.g., Ellis, 2010; Yu et al., 2019). The cognitive dimension pertains to how students pay attention to and process the received feedback, while the affective dimension includes emotional responses and reactions. The behavioral dimension involves whether and how students subsequently act upon the received feedback (Lipnevich & Smith, 2022; Zhang & Hyland, 2018). Consequently, the behavioral changes and developmental actions stemming from feedback can be situated within this latter dimension (Jellicoe & Forsythe, 2019). Moreover, Winstone and colleagues (2017) introduced the concept of proactive recipience of feedback, which refers to a state or activity in which learners actively participate in feedback processes and take co-responsibility for their effectiveness. The present study emphasizes the behavioral dimension of student feedback engagement (Ellis, 2010; Yu et al., 2019). More specifically, we consider two types of student feedback engagement within this dimension, corresponding to the proposed categorization of Handley and colleagues (2011). The first type, intentional feedback engagement, represents an individual's willingness to engage with feedback (i.e., readiness to invest time and effort). The second type, behavioral feedback engagement, reflects an individual’s active feedback engagement (i.e., taking actions) (Handley et al., 2011).

Reviews on educational feedback studies reveal that many of these studies focus on how feedback should be provided to effectively facilitate student outcomes (e.g., showing less favorable effects of feedback that relies on punishment/reward) (e.g., Van der Kleij et al., 2019). However, such investigations consequently view feedback as a more unidirectional and linear process driven by teachers, hereby neglecting student feedback engagement (Ajjawi & Boud, 2017; Van der Kleij & Lipnevich, 2020). In addition, findings across these feedback studies do not exhibit a coherent pattern (Shute, 2008; Winstone et al., 2017). Hence, researchers are shifting towards a socio-constructivist perspective (Dann, 2017; Winstone et al., 2019) by arguing that feedback only becomes genuinely effective for student outcomes when students behave as active and dialogic agents rather than as passive receivers during the feedback process (Carless & Boud, 2018; Winstone et al., 2019). The relationship between student feedback engagement and positive student outcomes is also empirically supported (e.g., Zhang & Hyland, 2018). Nonetheless, empirical studies focusing on student feedback engagement still lack adequate representation in literature (Lipnevich & Panadero, 2021; Van der Kleij et al., 2019).

In the present study, we examine student feedback engagement within the framework of the TPB. This theory suggests that an individual's behavior is mainly determined by their intention to perform that behavior (Ajzen, 1991, 2012), which is supported by meta-analyses (e.g., Riebl et al., 2015). We similarly hypothesize a positive effect of intentional feedback engagement on behavioral feedback engagement (H1) (see Figure 1).

Role of Student Characteristics

Recently, the field of feedback research is increasingly investigating how various feedback design variables (e.g., timing) can contribute to student feedback engagement (Jonsson, 2013; Van der Kleij et al., 2019). However, Panadero and Lipnevich (2022) reviewed the major developed descriptive feedback models across educational levels and state that the student, along with their individual characteristics, should assume a more central role in such models. Indeed, students' responses to feedback tend to vary based on their diverse individual characteristics (Lipnevich & Panadero, 2021; Van der Kleij et al., 2019). Further examination of student feedback engagement and the role of student characteristics is therefore highly encouraged (Panadero, 2023), and researchers are gradually venturing into this direction (Adams et al., 2020; Winstone et al., 2019). This approach can lead to a more effective alignment of feedback with students' educational needs, thereby enhancing their feedback engagement and learning outcomes (Panadero & Lipnevich, 2022; Winstone & Nash, 2023).

The present study incorporates the role of feedback self-efficacy as a student characteristic and its possible determinant effect on the association between feedback and student feedback engagement. This deliberate selection to focus on feedback self-efficacy stems from the decision to (partially) adopt the TPB as the theoretical framework for this study. Feedback self-efficacy, in fact, closely parallels the concept of perceived behavioral control within the TPB (Ajzen, 1985; Bosnjak et al., 2020).

Feedback Self-Efficacy

Self-efficacy indicates an individual's belief in their competences to perform a behavior (Bandura, 1977). In the case of feedback self-efficacy, this behavior refers to feedback engagement (Linderbaum & Levy, 2010; Winstone et al., 2019). Feedback self-efficacy is part of an individual’s feedback orientation (Linderbaum & Levy, 2010) and differs from academic self-efficacy, as the latter encompasses a broader belief in one’s abilities to engage in effective study behaviors (Sander & Sanders, 2009). Ajzen (1985) incorporated this self-efficacy construct into the Theory of Reasoned Action (Ajzen & Fishbein, 1980) in the form of perceived behavioral control. The addition of perceived behavioral control (hereafter: self-efficacy) as the third determinant of behavioral intention (alongside attitude towards the behavior and subjective norm) eventually led to the development of the TPB (Ajzen, 1991, 2012).

A review (Schneider & Preckel, 2017) and meta-analysis (Talsma et al., 2018) show a positive relationship between academic self-efficacy and academic achievement across various educational stages. Within the feedback context, researchers also demonstrate that academic self-efficacy mediates the relationship between feedback and academic achievement (e.g., Brown et al., 2016). Specifically focusing on student feedback engagement, studies indicate a positive association between academic self-efficacy and use of feedback (Adams et al., 2020; Handley et al., 2011), feedback self-efficacy and use of feedback (Winstone et al., 2019), and between feedback self-efficacy and readiness-to-engage (Handley et al., 2011). These findings are also supported by the TPB, as self-efficacy (i.e., feedback self-efficacy) influences behavioral intention (i.e., intentional feedback engagement) and actual behavior (i.e., behavioral feedback engagement). Indeed, individuals with higher feedback self-efficacy tend to exhibit an increased sense of control, and seem to be more confident in effectively managing the feedback they receive (Adams et al., 2020; Putwain et al., 2013; but see

Vancouver and Kendall (2006) for potential drawbacks of (very) high self-efficacy). Furthermore, behavioral intention mediates the relationship between self-efficacy and behavior (Ajzen, 1991, 2012), implying that individuals who believe in their ability to perform a behavior are more likely to perform that behavior when they have a strong intention to do so. Therefore, the present study expects to find a positive effect of feedback self-efficacy on intentional feedback engagement (H2) and behavioral feedback engagement (H3), and a mediation of intentional feedback engagement in the relationship between feedback self-efficacy and behavioral feedback engagement (H4) (see Figure 1).

Additionally, researchers highlight the potential role of student characteristics, such as feedback self-efficacy, as mediators in the relationship between feedback and student feedback engagement (Lipnevich & Panadero, 2021; Panadero, 2023). In other words, students' belief in their own feedback engagement capabilities (partially) explains the extent to which students engage with received feedback. Indeed, feedback self-efficacy can contribute to interpreting negative feedback as less threatening and more as valuable learning opportunities or challenges to overcome (Adams et al., 2020; Putwain et al., 2013). Moreover, the TPB emphasizes the importance of recognizing that self-efficacy (and the other antecedents of intention and behavior) are functions of underlying beliefs. For self-efficacy, this relates to one's control beliefs, which can be influenced by various background factors, categorized into individual (e.g., personality), social (e.g., education), and information-related (e.g., intervention) factors (Ajzen, 2020; Ajzen & Fishbein, 2005). Feedback literature similarly indicates the favorable impact of more positive, concrete, process-level etc. feedback on (academic) self-efficacy (Brown et al., 2016; Hattie & Timperley, 2007). However, TPB-based research often disregards this potential influence of background variables on the antecedents of intention. The present study includes the feedback received by students as a background factor. As such, we predict a positive effect of more favorable feedback on feedback self-efficacy (H5) and, consequently, a mediation of feedback self-efficacy in the relationship between feedback and intentional feedback engagement (H6) (see Figure 1).

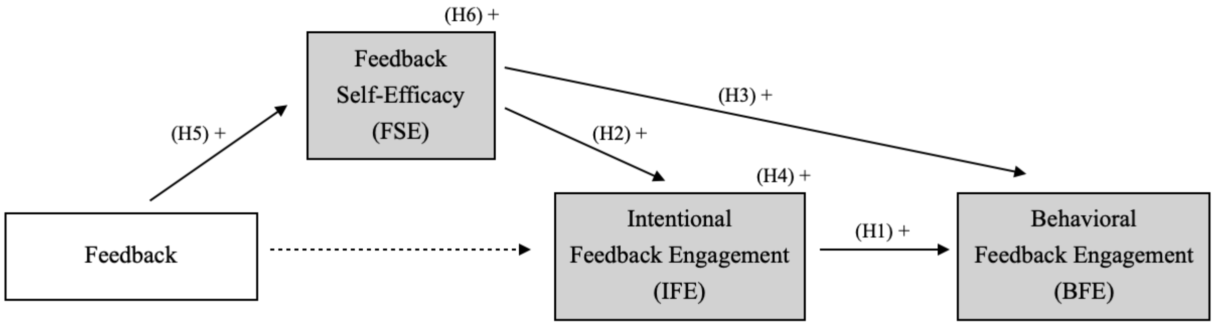
Present Study

The present study uses a TPB-based model for study (re)orientation and remediation advice in higher education. We aim to predict and explain student feedback engagement to develop targeted interventions that can facilitate feedback-engaging behavior (Panadero, 2023; Winstone & Nash, 2023), thereby enabling (at-risk) first-year university students to experience a successful study career. Our model's feedback refers to the feedback provided to these students at the onset of their higher education journey. This feedback includes their first-year validated chance of study success (i.e., (very) low, (fairly) high) prediction based on their background/(non-)cognitive competences, supplemented with recommendations for remediation/competence training activities.

Specifically, we hypothesize that intentional feedback engagement has a positive effect on behavioral feedback engagement (H1), and feedback self-efficacy similarly on intentional feedback engagement (H2) and behavioral feedback engagement (H3) (Ajzen, 1991, 2012). Here, feedback self-efficacy constitutes the equivalent of perceived behavioral control of the TPB (Ajzen, 1985; Bosnjak et al., 2020). Also, consideration is given to the mediation of intentional feedback engagement in the relationship between feedback self-efficacy and behavioral feedback engagement (H4) (Ajzen, 1991, 2012). Additionally, we anticipate that positive feedback increases feedback self-efficacy (H5) (Ajzen & Fishbein, 2005; Brown et al., 2016). Lastly, the mediation of feedback self-efficacy in the relationship between feedback and intentional feedback engagement is assumed (H6) (Lipnevich & Panadero, 2021). A graphical representation of our hypothesized model for the present study is displayed in Figure 1.

Figure 1

Hypothesized Model Using a Modified Version of the Theory of Planned Behavior



Note. The solid and dashed lines represent the hypothesized significant and non-significant paths, respectively. H4 displays the mediation of intentional feedback engagement between feedback self-efficacy and behavioral feedback engagement, and H6 the mediation of feedback self-efficacy between feedback and intentional feedback engagement. Feedback refers to the feedback first-year university students receive at the start of their study career regarding feasibility and recommendations for remediation/competence training initiatives.

Method

Participants

For the present study, we used data from a large Western European university that ranked in the top 75 of the Academic Ranking of World Universities (formerly Shanghai Ranking, see). The Ethics Committee at Ghent University, FPPW, granted a favorable recommendation for the project. Our sample comprises full-time first-year university students in an open access environment[1], excluding re-registrants, who have a uniform curriculum within a study program. Longitudinal data of N = 392 (62% female, 42% (very) low first-year predicted chance of study success) first-year university students in the academic years 2021-2022 and 2022-2023 were utilized, as these students were surveyed through feedback

[1] Besides completing secondary education, admission requirements only apply for Medicine, Dentistry, and Performing and Visual Arts, where candidates must pass an entrance exam.

questionnaires at two measurement points before the first exam period (i.e., October and December). These feedback questionnaires addressed the feedback the students received after participating in the longitudinal university-wide study (re)orientation and remediation project by filling out the SIMON test battery (Fonteyne, 2017; Fonteyne et al., 2017). For more detail about the sample, see Appendix 4A, Table A1.

Measures

Feedback

First-year university students using the SIMON tool receive personalized computer-generated feedback[2] upon completion (see also the introduction's feedback section). For concrete feedback examples, we refer to Appendix 4B.

SIMON estimates students' first-year chance of study success using recursive feature elimination and cross-validation. Background factors (e.g., secondary education degree) and (non-)cognitive characteristics (e.g., motivation, mathematical skills), recognized as significant predictors for academic achievement, are considered (Fonteyne et al., 2017). In the present study, we distinguished between two student groups as determined within the SIMON project: students who received feedback with a (very) low first-year predicted chance of study success (= 0) and those with a (fairly) high predicted chance of study success (= 1). Please see Fonteyne and colleagues (2017) for details on the operationalization of these predicted chances of study success.

Feedback Engagement

Intentional Feedback Engagement was measured through the behavioral and developmental change dimension of the Feedback in Learning Scale (FLS) (Jellicoe & Forsythe, 2019). The adapted questionnaire consisted of six items (e.g., "I will search for

[2] The computer-generated feedback is based on algorithms and predictive models that were conceptualized and operationalized by experts (Fonteyne, 2017; Fonteyne et al., 2017).

study guidance activities in line with competences described in my received SIMON feedback”) (M = 2.5, SD = 0.8, Cronbach’s α = .87). Students rated their item agreement on a Likert-scale from 1 (totally not agree) to 5 (totally agree). The individuals’ scale scores were determined by averaging their item scores. Intentional feedback engagement was surveyed in October for both academic years, after the students received their feedback through the SIMON project (Fonteyne, 2017; Fonteyne et al., 2017). The survey can be found in Appendix 4C, Table C1.

Behavioral Feedback Engagement was assessed through the same questionnaire employed for intentional feedback engagement (i.e., FLS; Jellicoe & Forsythe, 2019), with the verb form as the only difference (e.g., “I have searched for study guidance activities in line with competences described in my received SIMON feedback”) (M = 2.1, SD = 0.7, Cronbach’s α = .83). Again, the average of the item scores was used to determine the individuals’ scale scores. Behavioral feedback engagement was questioned in December for both academic years. The survey can be found in Appendix 4C, Table C2.

Feedback Self-Efficacy

Feedback Self-Efficacy was measured by using the same-named subscale of the Feedback Orientation Scale (FOS) (Linderbaum & Levy, 2010). The subscale consisted of five items (e.g., “I believe that I have the ability to deal with feedback effectively”) (M = 3.6, SD = 0.7, Cronbach’s α = .85). Students rated their item agreement on a 5-point Likert-scale. The individuals’ scale scores were obtained by averaging their item scores. Feedback self-efficacy was surveyed in October for both academic years, together with intentional feedback engagement.

Analyses

First, to have an overview of the data, we calculated Pearson correlations between the included continuous variables. For the correlations between the continuous variables and the dichotomous variable *Feedback*, point-biserial correlations were used.

Second, we tested our model with the predetermined set of hypotheses through a path analysis of a structural equation model using maximum likelihood estimation through the R package lavaan (Rosseel, 2012). The fit of our model was evaluated using different goodness-of-fit indices (i.e., chi-square test, Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR)).[3] To calculate better standard errors (and *p*-values) of the assumed indirect effects (i.e., mediations) compared to the standard approach (i.e., the delta or Sobel method), the bootstrap method was used (Bollen & Stine, 1992; Rosseel, 2012). We also added a fixed seed value for random sampling to ensure the results' reproducibility and comparability. Furthermore, we applied the Benjamini-Hochberg False Discovery Rate (FDR) procedure to adjust *p*-values for multiple testing, striking a balance between maximizing power and still effectively controlling Type I errors (Benjamini & Hochberg, 1995).

Results

Path Analysis

The goodness-of-fit indices indicate a good fit of the data to the model. The observed and model-implied covariance matrices can be found in Appendix 4D, Tables D1 and D2. Figure 2 displays the path analysis results, presenting the standardized regression coefficients with the FDR-adjusted significance levels, and the explained variances for the endogenous

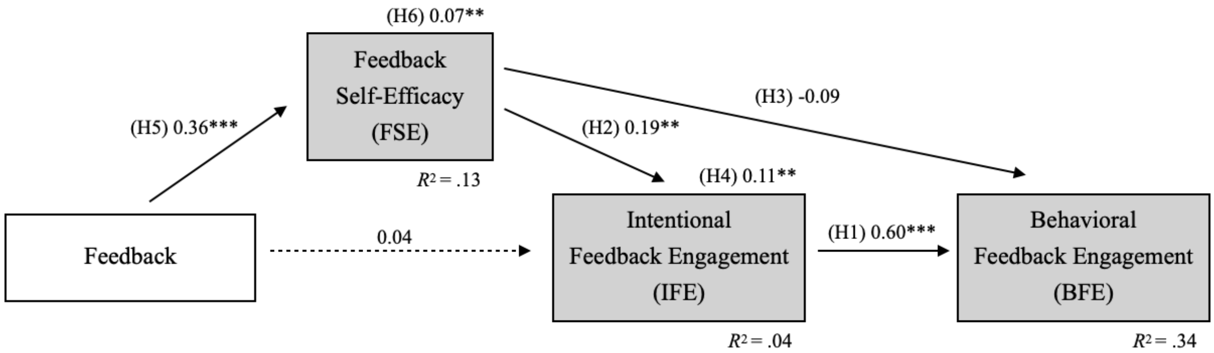
---

[3] Ideally, the chi-square test is not significant, incremental fit indices like CFI > .90, and measures such as RMSEA and SRMR < .06 for a good fit or between .06 and .08 for an acceptable fit (Rosseel, 2012).

variables within the model. For the statistics regarding the direct, indirect, and total effects within the path model, see Table 1.

Figure 2

Path Analysis Results (N = 392)



Note. ****p* < .001; ***p* < .01; **p* < .05. The *p*-values are FDR-adjusted using the Benjamini-Hochberg procedure for multiple testing. R^2 = explained variance. The solid and dashed lines represent the hypothesized significant and non-significant effects, respectively. The values associated with the lines are the standardized path coefficients. H4 displays the mediation of intentional feedback engagement between feedback self-efficacy and behavioral feedback engagement, and H6 the mediation of feedback self-efficacy between feedback and intentional feedback engagement. Model fit statistics: $\chi^2(1) = 1.57$, $p = .210$; CFI = 0.998; RMSEA = .038; SRMR = .015.

Hypotheses Testing

H1 stated that intentional feedback engagement has a positive effect on behavioral feedback engagement. We indeed find a significant effect of intentional feedback engagement on behavioral feedback engagement ($r = .58$). As the intentional feedback engagement level increases, so does the behavioral feedback engagement level. H1 is thus confirmed. Additionally, H2 assumed a positive effect of feedback self-efficacy on intentional feedback engagement. The results show that feedback self-efficacy significantly influences intentional feedback engagement ($r = .20$). Higher feedback self-efficacy is associated with higher intentional feedback engagement, confirming H2. Also, H3 stated that feedback self-efficacy has a positive effect on behavioral feedback engagement. We observe no significant direct effect between feedback self-efficacy and behavioral feedback engagement (r = .08), so H3 cannot be confirmed. Further, H4 proposed a mediation of intentional feedback engagement in the relationship between feedback self-efficacy and behavioral feedback engagement. The bootstrap method demonstrates a significant indirect effect of feedback self-efficacy on behavioral feedback engagement, fully mediated by intentional feedback engagement. Higher feedback self-efficacy increases the behavioral feedback engagement level, but this relationship completely relies on the intentional feedback engagement level (the higher intentional feedback engagement, the higher behavioral feedback engagement). As such, H4 is confirmed. Moreover, H5 stated a positive effect of feedback on feedback self-efficacy. We find a significant relationship between feedback and feedback self-efficacy (r = .36). Students who received feedback with a (fairly) high chance of study success show higher feedback self-efficacy (M = 3.9, SD = 0.7) compared with students who received a (very) low chance of study success (M = 3.4, SD = 0.8), confirming H5. Finally, H6 assumed a mediation of feedback self-efficacy in the relationship between feedback and intentional feedback engagement. The analyses show a significant indirect effect of feedback on intentional feedback engagement (r = .11), fully mediated by feedback self-efficacy. Students who received a (fairly) high chance of study success show higher intentional feedback engagement (M = 2.6, SD = 0.9) compared with students who received a (very) low chance of study success (M = 2.4, SD = 0.8), but this association completely depends on their feedback self-efficacy level (the higher feedback self-efficacy, the higher intentional feedback engagement). Hence, we confirm H6. The detailed statistics of the direct, indirect, and total effects within the path model can be found in Table 1. For the concrete feedback formulations, please refer to Appendix 4B.

Table 1

Direct, Indirect, and Total Effects Within the Path Model

	B	SE_B	*b*	*z*	p^1	CI	
						LL	UL
BFE ~ FSE + IFE							
Direct IFE (H1)	0.51	0.04	0.60	13.12	< .001***	0.44	0.59
Direct FSE (H3)	-0.08	0.04	-0.09	-2.01	.060	-0.17	0.00
Indirect FSE (H4)	0.11	0.03	0.11	3.29	.003**	0.04	0.17
Total FSE	0.03	0.05	0.03	0.54	.589	-0.08	0.12
IFE ~ Feedback + FSE							
Direct FSE (H2)	0.21	0.06	0.19	3.37	.003**	0.09	0.34
Direct feedback	0.07	0.09	0.04	0.77	.454	-0.11	0.24
Indirect feedback (H6)	0.12	0.04	0.07	3.06	.008**	0.05	0.20
Total feedback	0.19	0.09	0.11	2.15	.048*	0.02	0.36
FSE ~ Feedback							
Direct feedback (H5)	0.57	0.08	0.36	7.47	< .001***	0.42	0.72
BFE ~ Feedback + FSE							
Indirect feedback	-0.05	0.02	-0.03	-2.09	.057	-0.09	0.00
BFE ~ Feedback + IFE							
Indirect feedback	0.04	0.05	0.02	0.76	.454	-0.06	0.13
BFE ~ Feedback + FSE + IFE							
Indirect feedback	0.06	0.02	0.04	3.03	.008**	0.02	0.11

Note. ***$p < .001$, **$p < .01$, *$p < .05$. [1]FDR-adjusted using the Benjamini-Hochberg procedure for multiple testing. FSE = Feedback Self-Efficacy, IFE = Intentional Feedback Engagement, BFE = Behavioral Feedback Engagement. *B* = unstandardized path coefficients, *b* = standardized path coefficients, CI = bootstrapped 95% confidence intervals, LL = lower limit, UL = upper limit.

Discussion

In (open access) higher education, (first-year) fail and dropout rates are alarmingly high (OECD, 2022; Schelfhout et al., 2022). Hence, supporting students from the start of their higher education journey is likely to hold advantage for both students and society (OECD, 2022). Addressing this challenge ideally involves providing students with feedback on their validated first-year academic achievement prediction, derived from (non-)cognitive competencies, alongside recommendations for remediation and competence training activities. This feedback exceeds the conventional feedback typically associated with a specific competence (Kluger & DeNisi, 1996; Wisniewski et al., 2020).

Acknowledging the multidimensional nature of feedback is important, given its diverse forms that can yield distinct impacts on different student outcomes (Hattie & Timperley, 2007; Wisniewski et al., 2020). Previous feedback effectiveness studies mainly focused on identifying factors associated with feedback design and thus with feedback *provision.* However, such research neglects feedback *reception*, including students' active role in the feedback process (Boud & Molloy, 2013; Van der Kleij & Lipnevich, 2020). Meanwhile, feedback literature increasingly emphasizes how and to what extent students engage with their received feedback (i.e., socio-constructivist approach to feedback), which is essential for ultimately achieving improved student outcomes (Dann, 2017; Winstone et al., 2019). Nevertheless, theory-based empirical studies regarding student characteristics as possible determinants of student feedback engagement and its underlying mechanisms are limited, necessitating additional research to advance our comprehension and educational practices in this area (Panadero & Lipnevich, 2022; Winstone & Nash, 2023).

The present study therefore evaluated a model grounded in the Theory of Planned Behavior (TPB), well-known for predicting and explaining various behaviors (Ajzen, 1991, 2012), while also integrating insights from previous feedback research. To investigate determinants and underlying mechanisms of student feedback engagement within higher education, we used longitudinal data from $N = 392$ first-year university students and so differentiated between their intentional (October) and behavioral feedback engagement (December). Additionally, the proposed model incorporated students' feedback self-efficacy and the feedback they received at the start of their higher education journey regarding feasibility and recommendations for remediation/competence training initiatives.

Feedback Self-Efficacy, Intentional and Behavioral Feedback Engagement

Our findings confirm that students with higher intentional feedback engagement show higher behavioral feedback engagement, signifying more actual engagement with received feedback when the corresponding intention is more present, and thus consistent with the TPB's core idea that behavioral intention influences the actual behavior (Ajzen, 1991, 2012). This alignment is further supported by meta-analyses across various fields (e.g., Riebl et al., 2015). Indeed, intentions capture individuals' willingness to allocate effort and time towards performing a behavior (Ajzen, 1991; Handley et al., 2011).

Next, students with higher feedback self-efficacy exhibit higher intentional feedback engagement, reflecting a stronger readiness to invest effort and time to engage with received feedback when students have greater confidence in their feedback engagement capabilities. This finding confirms previous feedback research (Handley et al., 2011) and matches with one of the basic tenets of the TPB concerning perceived behavioral control (i.e., self-efficacy) that, among other things, determines behavioral intention (Ajzen, 1991, 2012). Individuals who consider themselves capable of succeeding in a task have more self-confidence and an increased sense of control compared with those who harbor uncertainty about their capabilities (Adams et al., 2020; Bandura, 1977). Linked to the Self-Determination Theory, experiencing a sense of competence and autonomy can indeed foster motivation (Deci & Ryan, 2008), which is assumed to be captured in behavioral intention (Ajzen, 1991, 2012).

Further, we do not find a direct effect of feedback self-efficacy on behavioral feedback engagement, but a full mediation of intentional feedback engagement. Students with higher feedback self-efficacy show higher behavioral feedback engagement, but only through their intentional feedback engagement level. This result implies that even if students believe they can effectively engage with feedback, they will only do so if they also intend to engage with the feedback. Meta-analyses including TPB-based research in non-educational areas do show that perceived behavioral control (i.e., self-efficacy) also directly influences behavior (e.g., Hirschey et al., 2020). In the educational context specifically, but not focusing on student feedback engagement, TPB-based studies indicate inconsistent results in this regard (e.g., Opoku et al., 2021). Additionally, Winstone and colleagues (2019) find an association between feedback self-efficacy and use of feedback. However, Ajzen (1991, 2012) highlights that the importance of intention and perceived behavioral control in predicting behavior can vary across situations and behaviors. One of these predictors may be more crucial, or even the sole determinant of behavior, depending on the context. In fact, not finding a direct effect of feedback self-efficacy on behavioral feedback engagement fits the TPB's predecessor, the Theory of Reasoned Action (Ajzen & Fishbein, 1980), which posits that individuals have voluntary control over behavior and thus results in the absence of perceived behavioral control in this theory.

Feedback (Self-Efficacy) and Intentional Feedback Engagement

The present study verifies our expectation that students who received feedback indicating a (fairly) high chance of study success demonstrate higher feedback self-efficacy compared with students who received feedback indicating a (very) low chance of study success. The TPB postulates that belief in one’s own competences is indeed a function of control beliefs, that can be influenced by information-related factors like feedback (Ajzen, 2020; Ajzen & Fishbein, 2005). Previous feedback studies also find that positive feedback is associated with higher (academic) self-efficacy (and conversely for negative feedback) (Brown et al., 2016; Peifer et al., 2020). This pattern is likely applicable to feedback self-efficacy as well. Positive/success versus negative/failure feedback can provide students with a sense of recognition and appreciation, which can boost their emotional positivity (Peifer et al., 2020; Winstone et al., 2017). Additionally, this uplift in emotional/physiological state is an important source of self-efficacy (Bandura, 1977).

Finally, upon closer investigation of the relationship between received feedback, feedback self-efficacy, and intentional feedback engagement, the results demonstrate a full mediation of feedback self-efficacy. Students who received feedback indicating a (fairly) high chance of study success (compared with a (very) low chance of study success) show higher intention to engage with the feedback, but this relationship depends on their belief in the own feedback engagement capabilities. Noteworthy, this finding implies that at-risk students are less reached compared with low-risk students due to negative received feedback reducing their feedback self-efficacy.[4] Additionally, this observed mediation confirms that student characteristics like feedback self-efficacy determine the extent to which students engage with received feedback (e.g., Lipnevich & Panadero, 2021). Students with greater feedback self-efficacy may perceive negative/failure feedback as less intimidating and more as valuable learning opportunities or challenges to conquer (Adams et al., 2020; Putwain et al., 2013). Enhancing students' responsiveness to feedback can thus be achieved by strengthening their feedback self-efficacy. Indeed, Warner and French (2020) refer to self-efficacy-based intervention studies across various domains with small-to-medium effect sizes for behavior. These interventions target at least one of the four established sources of self-efficacy (Bandura, 1977). For student feedback engagement specifically, researchers developed a confidence-building toolkit (Winstone et al., 2019). However, the toolkit's (quantitative) effectiveness on feedback self-efficacy remains unexplored in a sufficiently large sample.

---

[4] An extra linear regression on these students who received feedback indicating a (very) low chance of study success reveals a significant effect of feedback self-efficacy on intentional feedback engagement ($F(1, 163) = 7.58$, $p = .007$, $R^2 = .04$, $B = .23$). Within the group of at-risk students, those with higher feedback self-efficacy thus demonstrate higher intentional feedback engagement ($M = 2.5$, $SD = 0.9$) versus at-risk students with lower feedback self-efficacy ($M = 2.2$, $SD = 0.8$).

Overall, the present study's results show that the TPB provides a favorable theoretical framework in the feedback context, allowing for student characteristics to be considered in explaining student feedback engagement.

Strengths, Limitations and Future Research

The present study adopted a socio-constructivist approach to feedback, offering valuable insights into student feedback engagement (i.e., intentional and behavioral feedback engagement), its determinants and underlying mechanisms within higher education, based on a TPB-model. In addition to the favorable longitudinal data collection before the first exam period, the responses to the surveys result from self-report. An objective outcome measure, especially relevant for behavioral feedback engagement, is currently not present and recommended for future research. Additionally, the feedback in this study was computer-generated (see also Footnote 2), raising the question of potential disparities compared to human-generated feedback in students' responses. Nonetheless, meta-analytic evidence suggests that the quality and depth of feedback, rather than its source, are critical for improving learning outcomes (Van der Kleij et al., 2015). For student feedback engagement specifically, such empirical research is still understudied. Also, we did not use the full TBP as attitude and subjective norm were not included, which may explain the limited explained variance in intentional feedback engagement (4%). Noteworthy however, attitudes and norms may be more trait-like, making them less amenable to change. By contrast, perceived behavioral control (i.e., feedback self-efficacy) is more state-like (Bandura, 1977), rendering this determinant more practically relevant. Incorporating attitude and subjective norm in future research, along with feedback perceptions, may likely provide a more comprehensive understanding of the factors influencing intentional feedback engagement. Similarly, one could delve further into the underlying contributing beliefs for the antecedents of behavioral intention, as well as in the influencing background factors for these beliefs. Finally, we did not include (in)direct feedback loops between variables in the present study, as we relied on the TPB in which reciprocal relationships are not addressed. Nevertheless, understanding such feedback loops between various variables (e.g., direct feedback loop between feedback self-efficacy and intentional feedback engagement) is another interesting avenue for future exploration.

Conclusion

By using a TPB-based feedback model within a socio-constructivist framework, the present study underscores the importance of feedback self-efficacy on which interventions can focus to enhance student feedback engagement. Such insights are highly relevant in the (higher) education context, enabling students (at risk of failure) to receive more optimal support for the benefit of their study trajectory.

References

Adams, A.-M., Wilson, H., Money, J., Palmer-Conn, S., & Fearn, J. (2020). Student engagement with feedback and attainment: the role of academic self-efficacy. *Assessment & Evaluation in Higher Education*, *45*(2), 317–329. https://doi.org/10.1080/02602938.2019.1640184

Ajjawi, R., & Boud, D. (2017). Researching feedback dialogue: an interactional analysis approach. *Assessment & Evaluation in Higher Education*, *42*(2), 252–265. https://doi.org/10.1080/02602938.2015.1102863

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. . In J. Kuhl & J. Beckmann (Eds.), *Action-control: From cognition to behavior* (pp. 11–39). Springer.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Ajzen, I. (2012). The Theory of Planned Behavior. In *Handbook of Theories of Social Psychology: Volume 1* (pp. 438–459). SAGE Publications Ltd. https://doi.org/10.4135/9781446249215.n22

Ajzen, I. (2020). The theory of planned behavior: Frequently asked questions. *Human Behavior and Emerging Technologies*, *2*(4), 314–324. https://doi.org/10.1002/hbe2.195

Ajzen, I., & Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Behavior*. Prentice-Hall.

Ajzen, I., & Fishbein, M. (2005). The Influence of Attitudes on Behavior. In D. Albarracin, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173–221). Lawrence Erlbaum Associates.

Bandura, A. (1977). Self-efficacy: Toward a Unifying Theory of Behavioral Change. In *Psychological Review* (Vol. 84, Issue 2). https://search.proquest.com/docview/614275783?accountid=11077

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bosnjak, M., Ajzen, I., & Schmidt, P. (2020). The theory of planned behavior: Selected recent advances and applications. *Europe's Journal of Psychology*, *16*(3), 352–356. https://doi.org/10.5964/ejop.v16i3.3107

Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assessment & Evaluation in Higher Education*, *38*(6), 698–712. https://doi.org/10.1080/02602938.2012.691462

Brown, G. T. L., Peterson, E. R., & Yao, E. S. (2016). Student conceptions of feedback: Impact on self-regulation, self-efficacy, and academic achievement. *British Journal of Educational Psychology*, *86*(4), 606–629. https://doi.org/10.1111/bjep.12126

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education*, *31*(2), 219–233. https://doi.org/10.1080/03075070600572132

Carless, D. (2019). Feedback loops and the longer-term: towards feedback spirals. *Assessment & Evaluation in Higher Education*, *44*(5), 705–714. https://doi.org/10.1080/02602938.2018.1531108

Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Dann, R. (2017). *Developing feedback for pupil learning: Teaching, learning and assessment in schools.* Routledge.

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

Ellis, R. (2010). A Framework for Investigating Oral and Written Corrective Feedback. *Studies in Second Language Acquisition*, *32*(2), 335–349. https://doi.org/10.1017/S0272263109990544

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.

Fonteyne, L., Duyck, W., & de Fruyt, F. (2017). Program-specific prediction of academic achievement on the basis of cognitive and non-cognitive factors. *Learning and Individual Differences*, *56*, 34–48. https://doi.org/10.1016/j.lindif.2017.05.003

Handley, K., Price, M., & Millar, J. (2011). Beyond 'doing time': investigating the concept of student engagement with feedback. *Oxford Review of Education*, *37*(4), 543–560. https://doi.org/10.1080/03054985.2011.604951

Hattie, J., Gan, M., & Brooks, C. (2017). Instruction based on feedback. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 290–324). Routledge.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hirschey, R., Bryant, A. L., Macek, C., Battaglini, C., Santacroce, S., Courneya, K. S., Walker, J. S., Avishai, A., & Sheeran, P. (2020). Predicting physical activity among cancer survivors: Meta-analytic path modeling of longitudinal studies. *Health Psychology*, *39*(4), 269–280. https://doi.org/10.1037/hea0000845

Jellicoe, M., & Forsythe, A. (2019). The Development and Validation of the Feedback in Learning Scale (FLS). *Frontiers in Education*, *4*. https://doi.org/10.3389/feduc.2019.00084

Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education*, *14*(1), 63–76. https://doi.org/10.1177/1469787412467125

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284. https://doi.org/10.1037/0033-2909.119.2.254

Linderbaum, B. A., & Levy, P. E. (2010). The Development and Validation of the Feedback Orientation Scale (FOS). *Journal of Management*, *36*(6), 1372–1405. https://doi.org/10.1177/0149206310373145

Lipnevich, A. A., & Panadero, E. (2021). A Review of Feedback Models and Theories: Descriptions, Definitions, and Conclusions. *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.720195

Lipnevich, A. A., & Smith, J. K. (2022). Student – Feedback Interaction Model: Revised. *Studies in Educational Evaluation*, *75*, 101208. https://doi.org/10.1016/j.stueduc.2022.101208

Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: perceptions of students and academics. *Assessment & Evaluation in Higher Education*, *42*(2), 266–288. https://doi.org/10.1080/02602938.2015.1103365

OECD. (2022). *Education at a Glance 2022*. OECD. https://doi.org/10.1787/3197152b-en

Opoku, M. P., Cuskelly, M., Pedersen, S. J., & Rayner, C. S. (2021). Applying the theory of planned behaviour in assessments of teachers' intentions towards practicing inclusive education: a scoping review. *European Journal of Special Needs Education*, *36*(4), 577–592. https://doi.org/10.1080/08856257.2020.1779979

Panadero, E., & Lipnevich, A. A. (2022). A review of feedback models and typologies: Towards an integrative model of feedback elements. *Educational Research Review*, *35*, 100416. https://doi.org/10.1016/j.edurev.2021.100416

Peifer, C., Schönfeld, P., Wolters, G., Aust, F., & Margraf, J. (2020). Well Done! Effects of Positive Feedback on Perceived Self-Efficacy, Flow and Performance in a Mental Arithmetic Task. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01008

Putwain, D., Sander, P., & Larkin, D. (2013). Academic self-efficacy in study-related skills and behaviours: Relations with learning-related emotions and academic success. *British Journal of Educational Psychology*, *83*(4), 633–650. https://doi.org/10.1111/j.2044-8279.2012.02084.x

Riebl, S. K., Estabrooks, P. A., Dunsmore, J. C., Savla, J., Frisard, M. I., Dietrich, A. M., Peng, Y., Zhang, X., & Davy, B. M. (2015). A systematic literature review and meta-analysis: The Theory of Planned Behavior's application to understand and predict nutrition-related behaviors in youth. *Eating Behaviors*, *18*, 160–178. https://doi.org/10.1016/j.eatbeh.2015.05.016

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2). https://doi.org/10.18637/jss.v048.i02

Sander, P., & Sanders, L. (2009). Measuring academic behavioural confidence: the ABC scale revisited. *Studies in Higher Education*, *34*(1), 19–35. https://doi.org/10.1080/03075070802457058

Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., Derous, E., De Fruyt, F., & Duyck, W. (2022). How accurately do program-specific basic skills predict study success in open access higher education? *International Journal of Educational Research*, *111*, 101907. https://doi.org/10.1016/j.ijer.2021.101907

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, *143*(6), 565–600. https://doi.org/10.1037/bul0000098

Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Sobel, M. E. (1982). Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models. *Sociological Methodology*, *13*, 290. https://doi.org/10.2307/270723

Statistiek Vlaanderen. (2024, March 26). *Drop-out in het hoger onderwijs*. https://www.vlaanderen.be/statistiek-vlaanderen/onderwijs-en-vorming/drop-out-in-het-hoger-onderwijs

Talsma, K., Schüz, B., Schwarzer, R., & Norris, K. (2018). I believe, therefore I achieve (and vice versa): A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Learning and Individual Differences*, *61*, 136–150. https://doi.org/10.1016/j.lindif.2017.11.015

Van der Kleij, F. M., Adie, L. E., & Cumming, J. J. (2019). A meta-review of the student role in feedback. *International Journal of Educational Research*, *98*, 303–323. https://doi.org/10.1016/j.ijer.2019.09.005

Van der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of Feedback in a Computer-Based Learning Environment on Students' Learning Outcomes. *Review of Educational Research*, *85*(4), 475–511. https://doi.org/10.3102/0034654314564881

Van der Kleij, F. M., & Lipnevich, A. A. (2020). Student perceptions of assessment feedback: a critical scoping review and call for research. *Educational Assessment, Evaluation and Accountability*, *33*(2), 345–373. https://doi.org/10.1007/s11092-020-09331-x

Warner, L. M., & French, D. P. (2020). Self-Efficacy Interventions. In M. S. Hagger, L. D. Cameron, K. Hamilton, N. Hankonen, & T. Lintunen (Eds.), *The Handbook of Behavior Change* (pp. 461–477). Cambridge University Press.

Winstone, N. E., & Boud, D. (2022). The need to disentangle assessment and feedback in higher education. *Studies in Higher Education*, *47*(3), 656–667. https://doi.org/10.1080/03075079.2020.1779687

Winstone, N. E., Hepper, E. G., & Nash, R. A. (2019). Individual differences in self-reported use of assessment feedback: the mediating role of feedback beliefs. *Educational Psychology*, *41*(7), 844–862. https://doi.org/10.1080/01443410.2019.1693510

Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting Learners' Agentic Engagement With Feedback: A Systematic Review and a Taxonomy of

Recipience Processes. *Educational Psychologist*, *52*(1), 17–37. https://doi.org/10.1080/00461520.2016.1207538

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.03087

Yu, S., Zhang, Y., Zheng, Y., Yuan, K., & Zhang, L. (2019). Understanding student engagement with peer feedback on master's theses: a Macau study. *Assessment & Evaluation in Higher Education*, *44*(1), 50–65. https://doi.org/10.1080/02602938.2018.1467879

Zhang, Z. (Victor), & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, *36*, 90–102. https://doi.org/10.1016/j.asw.2018.02.004

Zimbardi, K., Colthorpe, K., Dekker, A., Engstrom, C., Bugarcic, A., Worthy, P., Victor, R., Chunduri, P., Lluka, L., & Long, P. (2017). Are they using my feedback? The extent of students' feedback use has a large impact on subsequent academic performance. *Assessment & Evaluation in Higher Education*, *42*(4), 625–644. https://doi.org/10.1080/02602938.2016.1174187

5

Expanding the Concept of Person-Environment Interest Fit: The Ins and Outs of Study Choice

Submitted to Journal of Counseling Psychology, 27th of June 2024.

Mona Bassleer, Stijn Schelfhout, Wouter Duyck & Nicolas Dirix.

Abstract

Person-environment (PE) interest fit between student and study program has recently received renewed attention because advanced regression-based methods show promising results for the prediction of academic achievement. Such methods also hold potential to enhance study counseling and reduce fail and dropout rates in higher education. Thus far, existing PE interest fit methods only acknowledge interest patterns that *align* with a specific environment, like a study program. During study orientation, however, students choose *between* different study programs, so that PE fit not only matters for what is chosen but also for what is not. The present study focused on study choice and introduced logistic regressed PE interest fit (LRIF), which differentiates between both interest patterns of students who choose a specific study program and those who choose another study program. Simultaneously, LRIF does not require additional data collection to predetermine environment interest profiles. We investigated the predictive validity of LRIF for study choice and compared it to traditional PE interest fit measures, including Euclidean distance (EDF) and correlation (CF) PE interest fit. Our analyses, involving N = 14,175 Flemish first-year university students across n = 31 independent study program subsets, demonstrate that LRIF accurately differentiates between students who choose a particular study program and those who choose another one. Moreover, LRIF rivals CF and outperforms EDF in predicting study choice. The LRIF method enriches the PE interest fit paradigm by encompassing both interest patterns that *align* with and *divert* from a specific environment. Practically, LRIF saves valuable time in counseling settings by eliminating the need for additional data collection to obtain environment interest profiles.

Introduction

Students tend to struggle during the transition from secondary to higher education, with first-year dropout rates ranging from 12% to 16% (OECD, 2022; Statistiek Vlaanderen, 2024) and fail rates reaching up to 60% in open access higher education contexts (Schelfhout, Wille, et al., 2022), incurring substantial costs for individuals and society (OECD, 2022). A better understanding of the congruence between students' and study programs' vocational interests can help students in making informed decisions about their academic path and increase the likelihood of success in their chosen field of study (Nauta, 2010; Schelfhout et al., 2021). Indeed, person-environment (PE) fit theories suggest that individuals strongly desire to match with their (work or study) environment (De Cooman et al., 2009; Holland, 1997; Oh et al., 2018), and that they are more likely to be successful in a fitting environment (e.g., Hoff et al., 2020; Lent et al., 1994; van Vianen, 2018). The impact of PE interest fit has been studied in both work (e.g., Nye et al., 2012, 2017) and higher education settings (e.g., Nye, Butt, et al., 2018; Schelfhout et al., 2022).

Meta-analyses on PE interest fit's predictive value for performance yield mixed results (Nye et al., 2012; Van Iddekinge et al., 2011). Inconsistent findings may stem from variations in the operationalization of PE interest fit (Nye, Prasad, et al., 2018; Xu & Li, 2020). However, literature applauds the growing shift in the way PE interest fit is measured over the years: From the use of congruence indices, that match the dominant interest dimension(s) between person and environment (Edwards, 1993; Tinsley, 2000), towards continuous and therefore more fine-grained PE interest fit methods (Tracey et al., 2012; Wille et al., 2014), and more recently also to a regression-based approach (Nye, Prasad, et al., 2018; Schelfhout, Bassleer, et al., 2022). Regressed PE interest fit methods provide a time-efficient advantage over the more traditional ones. Indeed, such methods have the important advantage that they do not need additional data collection from already enrolled individuals within an environment (e.g., third-year students within a study program) to predetermine environment interest profiles (see below) (Edwards, 1994; Xu & Li, 2020). Furthermore, studies using regressed PE interest fit typically show positive associations with work/study outcomes (Nye, Butt, et al., 2018; Nye, Prasad, et al., 2018; Schelfhout, Bassleer, et al., 2022). In the educational context specifically, regressed PE interest fit offers more accurate predictions for academic outcomes compared with traditional PE interest fit measures (Nye, Prasad, et al., 2018; Schelfhout, Bassleer, et al., 2022). This finding holds potential for enhancing study advice towards higher education (Nauta, 2010; Schelfhout et al., 2021) and improving fail/dropout rates (OECD, 2022; Su & Rounds, 2015) through the use of regressed PE interest fit. Considering current (regressed) PE interest fit literature, the present study identifies two issues that, to the best of our knowledge, need additional research to refine vocational interest literature and the use of PE interest fit in counseling contexts.

First, regressed PE interest fit is primarily used as means to investigate the prediction of academic achievement (Nye, Butt, et al., 2018; Nye et al., 2012; Schelfhout, Bassleer, et al., 2022). However, *study choice* in itself is at least as important as academic achievement in vocational interest literature, as vocational interests are arguably the best predictor of study choice (Rounds & Su, 2014; Stoll et al., 2017). The average correlation between a student's interest profile and their chosen study program over study programs can amount to $r = .70$ (Schelfhout et al., 2019). Researchers also state that study choice is a natural process in which future students explore different study programs, before making a study choice that matches their interests (Holland, 1997; Schneider, 1987).

Second, (regressed) PE interest fit is typically examined by matching an interest profile of an individual with interest patterns that align with a specific environment, such as a study program (Allen & Robbins, 2010; Tracey et al., 2012). Researchers do emphasize that individuals actively seek and select environments that fit with specific personal dispositions like their vocational interests, but that they also avoid environments that misfit (De Cooman & Vleugels, 2022; Holland, 1997; Schneider, 1987). Translated to the educational context, students choose to enroll in a particular study program (e.g., Psychology), implying by definition their choice to not enroll in other, competing study programs that may be less or more similar (e.g., Mathematics, Economics) (Feldman et al., 2001; Weidman, 2005). More specifically, students compare themselves with interest patterns that *align* with the Psychology study program for example, represented by students who choose this study program. At the same time, they compare themselves with interest patterns that *divert* from Psychology, represented by students who choose *another* study program (e.g., Mathematics, Economics). Thus, interest patterns of the study programs that students do not choose also have an impact on their study choice decision-making process. The existence of study programs with similar interest profiles will make study choice more difficult, irrespective of the PE interest fit between students and a given study program. By integrating the differentiation of interest patterns into a study's PE interest fit measure, attention is given to students' exploratory approach to study choice, enabling optimal support in this context.

The present study addresses both underexplored issues by introducing logistic regressed PE interest fit (LRIF). We first examine the validity of LRIF as an interest concept that predicts study choice. Second, we compare LRIF with traditional continuous PE interest fit measures in terms of their predictive validity for study choice. Opposed to traditional PE interest fit methods, LRIF does differentiate between interest patterns that align with and those that divert from a specific environment like a study program, while also eliminating the need for additional data collection to predetermine environment interest profiles. As such, the present study aims to contribute to vocational interest literature by focusing on regressed PE interest fit and study choice, with the ultimate goal of enhancing study advice for students during their study choice decision-making process and so improve the attrition rates in higher education.

Person–Environment Interest Fit with Interest Pattern Differentiation

Vocational interests refer to an individual's preferences for particular types of work-related activities and environments (Rounds & Su, 2014; Stoll et al., 2017). To categorize and understand an individual's vocational interests, Holland's (1997) RIASEC model is one of the most influential frameworks (Nauta, 2010). The circumplex model covers six vocational interest dimensions: Realistic (R), Investigative (I), Artistic (A), Social (S), Enterprising (E), and Conventional (C). By completing a RIASEC interest questionnaire, an individual's RIASEC profile can be assessed, providing scores for each of the six vocational interest dimensions (Holland, 1997; Schelfhout et al., 2019). Moreover, the dimensions of the RIASEC model allow to identify not only individuals' vocational interests but also those of environments.

Obtaining an environment's RIASEC profile can be achieved through various methods. The incumbent method is an appropriate illustration how interest patterns that align with an environment like a study program can be captured (Allen & Robbins, 2010; Schneider, 1987). By averaging out the scores of the students enrolled in a specific study program (i.e., incumbent students) on the six RIASEC dimensions, the importance of the single RIASEC dimensions as well as the relation between these dimensions can be estimated (i.e., the study program scores on each of the RIASEC dimensions and the intercorrelation of these dimensions, respectively). These interest patterns of the incumbent students are then used to establish the PE interest fit between a student and a study program. Note that this step of data collection becomes superfluous when using the LRIF method (see further).

PE fit is a prominent research construct in vocational psychology and organizational behavior as it is considered an important factor in understanding individual behavior and outcomes in various contexts (Guan et al., 2021; Jansen & Kristof-Brown, 2006). PE fit reflects the match/congruence between an individual's attributes (e.g., interests, values and abilities) and an environment (e.g., work, education, and leisure time), leading primarily to positive outcomes such as greater satisfaction (e.g., Hoff et al., 2020) and performance (e.g., Su & Nye, 2017). The present study specifically focuses on PE interest fit in the educational context, defined as the congruence between a student's vocational interests and the vocational study program environment (Hoff et al., 2020; Nye, Prasad, et al., 2018). Two traditional PE interest fit methods from literature are considered that compare the student to the study program interest profile and weigh differences on all six RIASEC dimensions: Euclidean distance and correlation.

Euclidean Distance and Correlation PE Interest Fit

First, Euclidean distance reduces the person and environment interest profiles into two points in Euclidean space (Wille et al., 2014). The computation of the distance between these points results in a continuous measure of PE interest fit (i.e., EDF) (Schelfhout, Bassleer, et al., 2022). Greater congruence is indicated by points that are situated closer together (Wilkins & Tracey, 2014). Second, the correlation method also provides a continuous PE interest fit measure (i.e., CF). This technique calculates the correlation between an individual's RIASEC scores and those of the environment (Allen & Robbins, 2010; Tracey et al., 2012). Researchers demonstrate that a higher degree of EDF or CF between a student's RIASEC profile and their chosen (STEM) study program interest profile increases the likelihood that the student ultimately select that study program (Schelfhout et al., 2019, 2021; Su & Rounds, 2015). The findings of a recent study consistently indicate that CF notably outperforms other PE interest fit measures (such as EDF and Angular Agreement) in predicting career outcomes (e.g., job and life satisfaction) (Xu & Li, 2020).

Though both techniques clearly have merit, they also by definition require availability of predetermined environment interest profiles, which are usually provided either by the incumbent method (Edwards, 1994; Nye et al., 2017) or through initiatives like O*Net (Rounds et al., 1999). More specifically, additional data collection is needed from already enrolled individuals in a specific environment, such as from third-year students within a study program. These data help to establish the study program interest profile, which is essential for calculating the PE interest fit between a (prospective) student and this study program (Edwards, 1994; Schelfhout et al., 2022). Environment interest profiles that are determined in such a way usually display average values for the RIASEC dimensions. Therefore, the variance within and the covariance between the RIASEC dimensions is not accounted for in these predetermined environment interest profiles, while researchers do emphasize the existence of interest (co)variance within an environment (Nye, Perlus, et al., 2018; Tracey et al., 2012). This limitation originated the use of other methods that consider more complex relationships, also including the variance within and covariance between interest dimensions across student incumbents, which put less constraints on the data (Edwards, 1993; van Vianen, 2018). One such method is regressed PE interest fit.

Regressed PE Interest Fit

Regressed PE interest fit calculates PE interest fit by regressing a criterion like academic achievement on the six RIASEC dimension variables (Edwards, 1994; Van Iddekinge et al., 2011). Researchers find that regressed PE interest fit offers a more precise evaluation of PE interest fit, along with improved predictive validity for academic achievement (Nye, Prasad, et al., 2018; Schelfhout, Bassleer, et al., 2022) and work satisfaction (Nye, Prasad, et al., 2018) when compared to traditional (non-)continuous PE interest fit measures. As such, using regressed PE interest fit allows for more accurate study guidance and support for students, which can then positively influence students' trajectories in higher education (Nauta, 2010; Su & Rounds, 2015). As a specific boon, regressed PE interest fit does not require predetermined environment interest profiles (Edwards, 1994; Nye et al., 2017). Consequently, collecting additional data is no longer necessary (Edwards, 1994; Schelfhout et al., 2022). In fact, regressed PE interest fit can generate the environment interest profiles itself, taking into account the variance within and covariance between interest dimensions: The regression coefficients, derived from the regression using individuals' RIASEC scores as predictors and a criterion such as study choice as the dependent variable, represent the environment interest profile (Schelfhout, Bassleer, et al., 2022). Despite the advantages, regressed PE interest fit has not been used to determine PE interest fit in function of study choice, nor has this method considered both interest patterns of students who choose a specific study program and of those who choose another study program.

Logistic Regressed PE Interest Fit. To differentiate between interest patterns that align with a study program like Psychology and interest patterns that divert from this study program, we propose LRIF that uses the study choice of students as the criterion. As for the more concrete operationalization of LRIF, students that choose Psychology for instance represent the interest patterns that align with the Psychology study program. Students that do not choose Psychology but another study program represent the interest patterns that divert from this Psychology study program (e.g., Mathematics, Economics). More formally and as shown in Eq. (1), the differentiation of both types of interest patterns is represented by a logistic regression and ultimately by the set of regression weights:

$$P(S) = \frac{e^{\beta_0+\beta_1 R+ \beta_2 I+ \beta_3 A+ \beta_4 S+ \beta_5 E+ \beta_6 C}}{1+ e^{\beta_0+\beta_1 R+ \beta_2 I+ \beta_3 A+ \beta_4 S+ \beta_5 E+ \beta_6 C}}, \tag{1}$$

with *S* being a binary variable that can take the value 1 (i.e., choosing the study program) or 0 (choosing another study program), *P(S)* returning the probability of choosing the study program, and with *R* to *C* representing the student scores on the RIASEC dimensions. β_0 to β_6 are regression coefficients and need to be estimated. Note that the expression in the exponent of Euler's number is linear in its parameters (Edwards, 1993; Nye, Prasad, et al., 2018). Also, LRIF only uses individual linear RIASEC terms in the exponent of the logistic regression. As such, the method reduces predictors, minimizing type I errors associated with an increasing number of predictors (Su et al., 2019). In fact, the entire logistic regression can also be written as a linear expression by using a logit transformation of the odds ratio as displayed in Eq. (2):

$$\text{logit}(P) = \ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 R + \beta_2 I + \beta_3 A + \beta_4 S + \beta_5 E + \beta_6 C, \qquad (2)$$

with *P* again returning the probability of choosing the study program, 1 – *P* returning the probability of choosing another study program, and with *R* to *C* representing the student scores on the RIASEC dimensions.

In sum, the present study contributes to researchers' call for increased use of more advanced regression-based methods to determine PE interest fit (Nye, Prasad, et al., 2018; Schelfhout, Bassleer, et al., 2022). Two underexplored issues in (regressed) PE interest fit literature are tackled by emphasizing study choice instead of academic achievement, and by incorporating the differentiation of interest patterns. Hence, the present study examines to what extent LRIF predicts study choice and thus can distinguish between students that choose a specific study program versus those who choose another study program. To further evaluate and refine this predictive validity of LRIF for study choice, we also compare the predictive validity for study choice offered by LRIF versus by EDF and CF. Unlike these latter two traditional continuous PE interest fit methods, LRIF has the conceptual advantage of differentiating between interest patterns that align with and those that divert from a specific environment like a study program. Moreover, LRIF has the practical advantage of not requiring additional data collection to predetermine environment interest profiles. In the educational context, regressed PE interest fit shows to perform better for predicting academic achievement when compared with congruence indices (Nye, Prasad, et al., 2018) and traditional continuous PE interest fit measures (Schelfhout, Bassleer, et al., 2022). However, these studies focus on academic achievement and the prediction of study choice by PE interest fit is not yet compared between methods. Moreover, research already reveals a high average correlation across study programs between student and study program interest profile (r = .70) (Schelfhout et al., 2019). We therefore expect an at least equally predictive validity for study choice of LRIF compared with CF and EDF, keeping in mind that LRIF offers conceptual and practical benefits.

Method

Participants

The present study analyzed data from a Western European university ranked in the top 75 in the ARWU ranking (formerly Shanghai Ranking, see https://www.shanghairanking.com/rankings/arwu/2022). The university has 11 faculties including 42 bachelor's programs, and utilizes a low private cost (i.e., annual tuition fee of about $1,075.74 or €979.60) open access system.[1] Full-time first-year students at this university follow a uniform curriculum within their chosen study program. Consequently, the study program environments of the present study are well suited to ascertain the pure influence of vocational interests on study choice as the biases of ability-based selection and costs are minimal.

At the start of each academic year, first-year higher education students are strongly encouraged to complete an online RIASEC questionnaire that identifies their vocational

[1] Except for the study programs Medicine, Dentistry, and Performing and Visual Arts, where students are required to pass an entrance examination, secondary education qualifications are sufficient for admission into a higher education study program.

interests. This questionnaire is part of a longitudinal university-wide study (re)orientation and remediation project (the SIMON project) (Fonteyne, 2017; Fonteyne et al., 2017). Only the study programs with a sample size of $N \geq 110$ first-year students (no re-registrants) were included. We like to refer to *Analyses* for more in-depth information about this methodological decision.

Eventually, we obtained data on students' vocational interests and study choice (i.e., students' chosen study program) from a large overall sample size of $N = 14{,}175$ (40% male, $M_{age} = 18$). This sample encompassed the three consecutive academic years 2016-2017, 2017-2018 and 2018-2019, and was distributed across 31 study programs. For more detail, see Appendix 5A, Table A1.

RIASEC Dimensions of Vocational Interests

The dimensions of the RIASEC model (Holland, 1997) were assessed using the SIMON-I instrument (Fonteyne, 2017; Fonteyne et al., 2017), which is designed to support students in the study (re)orientation process towards/in higher education. The SIMON-I instrument includes 173 items across six dimensions that ask students whether they would be interested in engaging in specific activities or pursuing particular occupations. The realistic dimension (R) was measured through 27 items (e.g., developing electronic systems, pilot; $\alpha = .93$) and the investigative dimension (I) through 33 items (e.g., analyzing statistics, researcher; $\alpha = .88$). The artistic dimension (A) was assessed through 30 items (e.g., designing webpages, photographer; $\alpha = .92$) and the social dimension (S) through 32 items (e.g., giving travel advice, teacher; $\alpha = .92$). The enterprising dimension (E) was assessed through 26 items (e.g., organizing a conference, lawyer; $\alpha = .93$) and the conventional dimension through 25 items (e.g., calculating prices, judge; $\alpha = .91$). To calculate the final student score on each dimension, the number of "yes" answers for that dimension was summed and divided by the total number of items for that dimension. This quotient was then multiplied by 100 to obtain a score between 0 and 100. The higher the score on a RIASEC dimension, the stronger the preference for that area of interest. The RIASEC dimensions were used as the independent variables in the LRIF method.

Study Choice

Before students could complete the RIASEC questionnaire, they had to indicate their chosen study program in their first year of higher education (therefore, we also knew students' non-chosen study programs). Study choice is thus binary (1 = choosing the study program, 0 = choosing another study program) and concerns the dependent variable in the LRIF method.

Logistic Regressed Person-Environment Interest Fit

Logistic regressed PE interest fit (LRIF) refers to the congruence between a student's RIASEC scores and a study program interest profile in higher education using logistic regression (linear in its parameters). To calculate LRIF for a specific study program environment, we complemented the incumbent method (Allen & Robbins, 2010). Indeed, both students who choose a particular study program as well as students who choose another study program were included. The study choice data of students and (randomly selected) non-students for a specific study program environment were thus binary logistic regressed on the (non-)students' RIASEC scores. As such, the regression coefficients were estimated for the study program RIASEC terms, representing the study program interest profile. The LRIF for a (non-)student in a specific study program was calculated by initially multiplying the student RIASEC term scores with the corresponding coefficients from the program-specific logistic regression, after which the six products were summed. This procedure resulted in the students' logit scores, which were then converted to probability scores. As such, combining Eqs. (1) and (2) from the introduction results in Eq. (3):

$$P(S) = \frac{e^{\text{logit}}}{1+e^{\text{logit}}} = \frac{e^{\beta_0+\beta_1 R+\beta_2 I+\beta_3 A+\beta_4 S+\beta_5 E+\beta_6 C}}{1+e^{\beta_0+\beta_1 R+\beta_2 I+\beta_3 A+\beta_4 S+\beta_5 E+\beta_6 C}}, \quad (3)$$

with *S* being a binary variable that can take the value 1 (i.e., choosing the study program) or 0 (i.e., choosing another study program), *P(S)* returning the probability of choosing the study program, and with *R* to *C* representing the students' scores on the RIASEC dimensions. β_0 to β_6 are the estimated regression coefficients. Hence, these probability scores represent the likelihood of choosing that particular study program for a student based on their RIASEC scores. This procedure was repeated for each of the 31 study programs included in the present study.

Analyses

In order to calculate (non-)students' LRIF, 31 data subsets were created. Each data subset included data from two first-year university student groups: students who choose a specific study program and an equal number of students, obtained through random sampling, who choose another study program. We decided to create a data subset only for those study programs where the group of first-year university students who choose the study program consisted of at least 110 individuals. To determine this minimum required sample size, we used a conservative approach by taking 50 individuals as the baseline. In doing so, study programs with an insubstantial number of students were eliminated. Also, we followed Peduzzi and colleagues (1996) who proposed to incorporate ten participants per included predictor in a model. Since we included the six RIASEC dimensions as predictors in our models, at least 110 students that choose a specific study program were required to integrate the study program in the present study. Additionally, we used the LRIF method with study choice as dependent variable and the (non-)students' RIASEC scores as predictors, which eventually resulted in 31 LRIF study program interest profiles via the estimated regression coefficients, and (non-)students' LRIF.

To first answer to what extent LRIF can predict study choice, 31 logistic regressions were performed with study choice as the dependent variable and LRIF as the predictor. We observed the model summaries, including Nagelkerke's pseudo R-squared (R^2_{Nag}; goodness of fit measure) (Nagelkerke, 1991), and the classification tables in order to show the models' fit. The closer R^2_{Nag} is to one, the better the fit of the model (Hosmer & Lemeshow, 2000; Nagelkerke, 1991). Moreover, ROC (Receiver Operating Characteristic) curves were developed to show the degree of effectiveness of our LRIF models in accurately classifying students into their chosen study program, and differentiating them from those students who choose another study program. For this purpose, we used the associated Area Under the ROC Curve (AUC) values, where $.70 < AUC \leq .80$ is acceptable, $.80 < AUC \leq .90$ excellent, and $AUC > .90$ is considered outstanding (Hosmer & Lemeshow, 2000). In addition, to give an indication of how large the difference is between the mean LRIF of the students who choose a particular study program versus those who choose another study program, the mean LRIF logit scores for both groups of students were calculated per study program. A paired samples *t*-test was performed on these 31 pairs of mean LRIF values, followed by a calculation of the effect size (i.e., Cohen's *d*). The following rule of thumb was applied, introduced by Sawilowsky (2009): $d = 0.01$ (very small), $d = 0.20$ (small), $d = 0.50$ (medium), $d = 0.80$ (large), $d = 1.20$ (very large) and $d = 2.00$ (huge).

Second, to compare the predictive validity for study choice offered by LRIF versus by EDF and CF, we evaluated the explained variance (i.e., R^2_{Nag})[2] in study choice by PE interest fit, established through logistic regression versus through Euclidean distance and correlation. For these last two methods, study program interest profiles had to be predetermined.

[2] Technically, explained variance does not apply to logistic regression. Instead, pseudo R-squared measures, like Nagelkerke's pseudo R-squared, evaluate model performance by comparing the deviance of the null model to that of the model with predictors (Nagelkerke, 1991).

Therefore, we implemented the incumbent method by using a separate dataset with the RIASEC scores of university students in their third year of a specific study program (academic years 2016-2017 to 2018-2019 were included). Specifically, the average RIASEC scores of these third-year university students in a particular study program ultimately formed the study program interest profile for that specific study program. An overview of these study program interest profiles and the corresponding sample sizes can be found in Appendix 5A, Table A2. Subsequently, for the calculation of EDF and CF, the Euclidean distance[3] and correlation between a first-year student's RIASEC scores and the predetermined RIASEC profile of a specific study program were calculated, respectively. Study choice was then binary logistic regressed on EDF and CF separately, to primarily gain insight into R^2_{Nag}. Again, this procedure was carried out for each of the 31 study programs included in the study. Additionally, the R^2_{Nag} values generated through the three different methods were compared using paired samples *t*-tests, for which Cohen's *d* effect sizes were calculated. Ultimately, we computed the correlations between the different PE interest fit measures, as well as the correlations between the LRIF and predetermined EDF/CF study program interest profiles to assess their correspondence.

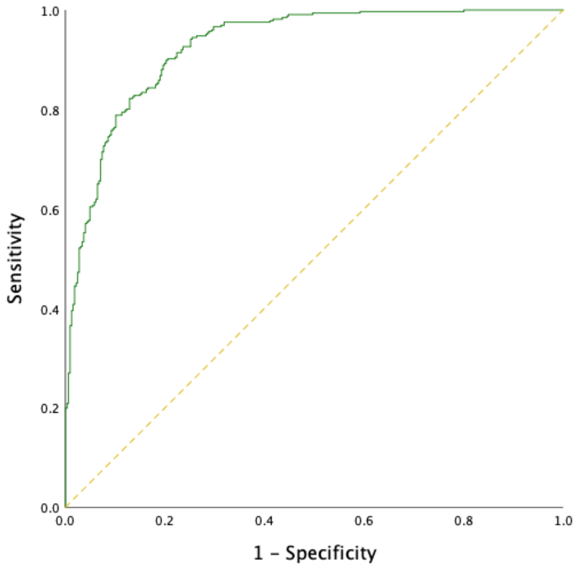
[3] The Euclidean distance was computed following the methodology described by Wille and colleagues (2014). For each student, two points were derived in Euclidean space, each consisting of two coordinates. The people-things (P/T) axis spans from point S to point R on the RIASEC hexagon. The P/T coordinate for each student and program was determined using the formula: $\frac{P}{T} = 2\ \times R + I - A - 2 \times S - E + C$. Similarly, the data-ideas (D/I) axis extends from a position between the E and C dimensions to a position between the A and S dimensions on the RIASEC hexagon. The D/I coordinate for each student and program was calculated using the formula: $\frac{D}{I} = 1.73 \times E + 1.73 \times C - 1.73 \times I - 1.73 \times A$. Finally, the Euclidean distance (ED) was computed for each student using the following formula: $\sqrt{\left(\left(\text{student}\,\frac{\text{P}}{\text{T}} - \text{program}\,\frac{P}{T}\right)^2 + \left(\text{student}\,\frac{\text{D}}{\text{I}} - \text{program}\,\frac{D}{I}\right)^2\right)}$.

Results

First, we evaluated the extent to which LRIF predicts study choice by conducting 31 logistic regressions. The pseudo R-squared values resulting from these analyses range from $R^2_{Nag} = .19$ to $R^2_{Nag} = .88$ (*M* = .49, *SD* = .17, 95% *CI* [.43, .54]). For example, the LRIF model for Communication Sciences manages to both correctly classify 85% of the students who choose Communication Sciences (i.e., sensitivity) and 82% of the students who choose another study program (i.e., specificity). Further, to test the ability of the LRIF models to distinguish between students that choose a particular study program versus those that choose another study program, we looked at the balance between sensitivity and specificity of our LRIF models by using ROC curves and the associated AUC values. The analyses show that the 31 AUC values range from *AUC* = .71 to *AUC* = .98 (*M* = .85, *SD* = .07, 95% *CI* [.83, .87]). Figure 1 presents the ROC curve for the study program Communication Sciences. For this study program, an AUC value of *AUC* = .93 is found with an asymptotic 95% *CI* [.91, .95]. The detailed results per study program are shown in Table 1. For the ROC curves per study program, see Appendix 5B, Figure B1. When generalizing our findings, a paired samples *t*-test ($t(30) = 8.77$, $p < .001$, $d = 1.57$) shows a significant higher mean PE interest fit (based on the logit scores) for students who choose a certain study program (*M* = 1.18, *SD* = 0.74, 95% *CI* [0.92, 1.44]) compared with students who choose another study program (*M* = -1.45, *SD* = 0.93, 95% *CI* [-1.77, -1.12]).

Figure 1

ROC Curve Study Program Communication Sciences



Note. Sensitivity refers to the proportion of true positives, or the number of students correctly classified as Communication Sciences students. 1 - Specificity refers to the proportion of false positives, or the number of students incorrectly identified as Communication Sciences students. The green ROC curve delimits The Area Under the Curve (AUC). AUC is a measure of the model's ability to distinguish between students who choose Communication Sciences and students who choose another study program, in this case based on their PE interest fit (i.e., LRIF). The AUC value for this model is 93%. The yellow dotted reference line refers to the 50% probability level benchmark of distinction.

Second, we compared the predictive validity for study choice by PE interest fit, established through LRIF versus through EDF and CF. The paired samples *t*-tests reveal a significant higher mean pseudo R-squared for LRIF ($M = .49$, $SD = .17$, 95% CI [.43, .54]) compared with EDF ($M = .28$, $SD = .11$, 95% CI [.24, .32]) ($t(30) = 7.85$, $p < .001$, $d = 1.43$). A similar result is found when comparing CF ($M = .46$, $SD = .17$, 95% CI [.40, .52]) with EDF, in favor of CF ($t(30) = 8.38$, $p < .001$, $d = 1.50$). The mean pseudo R-squared for LRIF is higher than for CF, but this difference is non-significant ($t(30) = 1.64$, $p = .112$, $d = 0.03$).

The correlation between the PE interest fit measures LRIF and CF amounts to $r = .68$, between LRIF and EDF to $r = -.18$, and between EDF and CF to $r = -.54$. Additionally, the correlations between the LRIF and predetermined EDF/CF study program interest profiles range from $r = .61$ to $r = .98$ ($M = .83$, $SD = .10$, 95% CI [.80, .87]). For the more extensive output per study program, we like to refer to Table 1.

Table 1

LRIF Study Program Interest Profiles, Goodness of Fit Measures per PE Interest Fit Measure and Bivariate Correlations LRIF and EDF/CF Study Program Interest Profiles

Study Program	LRIF Study Program Interest Profile	R^2_{LRIF}	R^2_{EDF}	R^2_{CF}	AUC_{LRIF}	r_{profiles}
Applied Language Studies	$-.175 - .698 \times R - 1.135 \times I + 1.138 \times A - .059 \times S - .126 \times E - .302 \times C$	.51	.29	.47	.86	.80
Art History	$-.537 - 1.416 \times R - .608 \times I + 3.9 \times A - .71 \times S - .102 \times E - 1.769 \times C$	.88	.34	.87	.98	.98
Bio Sciences	$-.058 + .384 \times R + .958 \times I - .651 \times A - .848 \times S - .246 \times E + .169 \times C$	.44	.23	.41	.84	.81
Biochemistry and Biotechnology	$-.104 - .543 \times R + 1.701 \times I + .029 \times A - .475 \times S - .873 \times E + .254 \times C$	.47	.29	.46	.85	.89
Biology	$-.228 - .253 \times R + 1.09 \times I + .264 \times A - .485 \times S - .862 \times E - .692 \times C$	.44	.39	.48	.83	.94
Biomedical Sciences	$-.079 - .628 \times R + 1.559 \times I - .406 \times A - .082 \times S - .292 \times E - .251 \times C$	.43	.24	.44	.84	.96
Bioscience Engineering	$-.08 + .091 \times R + 1.759 \times I - .421 \times A - .559 \times S + .084 \times E - .374 \times C$	.50	.19	.47	.86	.96
Business Administration	$-.127 - .198 \times R - 1.054 \times I - .087 \times A - .425 \times S + 1.248 \times E + .956 \times C$	.65	.59	.66	.92	.92
Communication Sciences	$-.214 - .728 \times R - .805 \times I + 1.696 \times A - .414 \times S + 1.59 \times E - .919 \times C$	.68	.28	.68	.93	.91
Computer Sciences	$-.201 + .754 \times R - .33 \times I + .037 \times A - 1.301 \times S - .345 \times E + .074 \times C$	.45	.35	.41	.86	.68
Criminological Sciences	$-.038 - .519 \times R - .298 \times I - .093 \times A + .54 \times S - .469 \times E + .392 \times C$	.20	.22	.22	.72	.61
(Applied/Business) Economics	$-.096 - .119 \times R - .496 \times I - .107 \times A - .856 \times S + 1.065 \times E + 1.264 \times C$	.64	.52	.63	.91	.90
Educational Sciences	$-.507 - .407 \times R - .931 \times I - .225 \times A + 2.904 \times S - .585 \times E - .044 \times C$	.73	.52	.68	.94	.88
Engineering	$-.162 + 1.533 \times R + .467 \times I - .573 \times A - 1.139 \times S + .372 \times E - .454 \times C$	.60	.36	.64	.90	.98
Engineering - Architecture	$-.233 + 1.911 \times R - .903 \times I + 1.946 \times A - 1.22 \times S - .193 \times E - .203 \times C$	.73	.12	.56	.94	.77
Engineering Technology	$-.014 + 2.912 \times R - .68 \times I - .251 \times A - .719 \times S - .267 \times E - .174 \times C$	.76	.36	.74	.95	.86
History	$-.01 - .711 \times R + .11 \times I + 1.1 \times A - .52 \times S + .097 \times E - .28 \times C$	.29	.10	.27	.79	.70
Law	$-.067 - 1.047 \times R - .268 \times I + .332 \times A - .182 \times S + .636 \times E + .642 \times C$	.38	.25	.41	.81	.80
Linguistics and Literature	$-.157 - .948 \times R - .588 \times I + 1.825 \times A - .159 \times S - .592 \times E - .303 \times C$	.59	.29	.55	.90	.79
Medicine	$-.051 - .086 \times R + .949 \times I - .512 \times A + .612 \times S - .339 \times E - .389 \times C$	.35	.16	.28	.80	.81
Oriental Languages and Cultures	$-.043 - .518 \times R - .336 \times I + .858 \times A - .276 \times S - .285 \times E - .191 \times C$	.28	.21	.29	.76	.67
Pharmaceutical Sciences	$-.049 - .623 \times R + 1.207 \times I - .344 \times A - .064 \times S - .535 \times E + .164 \times C$	.36	.23	.38	.80	.85

Study Program	LRIF Study Program Interest Profile	R^2_{LRIF}	R^2_{EDF}	R^2_{CF}	AUC_{LRIF}	$r_{profiles}$
Physical Education and Movement Sciences	$-.034 + .055 \times R - .022 \times I - .468 \times A + .426 \times S + .003 \times E - .749 \times C$	.19	.14	.08	.71	.64
Physical Therapy and Motor Rehabilitation	$-.059 + .008 \times R + .218 \times I - .64 \times A + .921 \times S - .675 \times E - .277 \times C$	.33	.25	.35	.79	.79
Physics and Astronomy	$-.134 + .307 \times R + 1.143 \times I + .423 \times S - 1.648 \times E - .312 \times O - .25 \times C$	.51	.30	.52	.87	.78
Political Sciences	$-.036 - .776 \times R - .258 \times I + .31 \times A - .248 \times S + 1.228 \times E - .602 \times C$	.34	.21	.35	.80	.84
Psychology	$-.113 - .524 \times R - .151 \times I - .005 \times A + 1.215 \times S - .185 \times E - .547 \times C$	.42	.32	.45	.83	.95
Public Administration and Management	$-.158 - .326 \times R - 1.098 \times I + .017 \times A + .34 \times S + 1.216 \times E + .487 \times C$	.55	.32	.51	.88	.86
Sociology	$-.133 - 1.213 \times R - .094 \times I + .78 \times A + .186 \times S + 1.047 \times E - 1.056 \times C$	.45	.21	.41	.84	.81
Speech Language and Hearing Sciences	$-.198 - .498 \times R - .232 \times I - .082 \times A + 1.826 \times S - .744 \times E - .173 \times C$	.56	.33	.58	.88	.89
Veterinary Medicine	$-.044 - .309 \times R + 1.115 \times I - .25 \times A - .219 \times S - .618 \times E - .05 \times C$	.33	.22	.29	.79	.78

Note. R = realistic interest dimension, I = investigative interest dimension, A = artistic interest dimension, S = social interest dimension, E = enterprising interest dimension, C = conventional interest dimension. R^2 = explained population variance, measured through Nagelkerke's pseudo R-squared (Nagelkerke, 1991) (see also Footnote 2). AUC = Area Under the Curve. LRIF = logistic regressed PE interest fit, CF = correlation PE interest fit, EDF = Euclidean distance PE interest fit. The LRIF study program interest profiles are the result of the logistic regressions of study choice on the (non-)students' RIASEC scores. The Nagelkerke's pseudo R-squared values were obtained from the logistic regressions of study choice on LRIF/EDF/CF. The AUC values indicate how well the LRIF models can distinguish between students that choose a certain study program versus those that choose another study program. The last column presents the correlations between the predetermined study program interest profiles, on which CF and EDF were based, and the LRIF study program interest profiles.

Discussion

Students face a challenging transition from secondary to higher education, which is reflected in high first-year fail and dropout rates, especially in higher education systems without strict prior admission procedures (OECD, 2022; Schelfhout, Wille, et al., 2022). Consequently, counseling based on well-informed and reliable study advice is essential. Previous research highlights the critical role of vocational interests in shaping educational and occupational choices (Rounds & Su, 2014; Stoll et al., 2017). More specifically, person-environment (PE) interest fit between the vocational interests of an individual and an environment like a study program has recently received renewed attention in literature. Here, regressed PE interest fit measures overcome the limitations of traditional congruence indices (Edwards, 1993; Tinsley, 2000), and render promising results towards predicting work and study outcomes (Nye, Butt, et al., 2018; Nye, Prasad, et al., 2018; Schelfhout, Bassleer, et al., 2022). Moreover, regressed PE interest fit provides a time-efficient advantage over traditional (continuous) PE interest fit methods. This method indeed eliminates the need for additional data collection to predetermine environment interest profiles (Edwards, 1994; Schelfhout et al., 2022), while also considering the variance within and covariance between interest dimensions (Tracey et al., 2012; Xu & Li, 2020).

The present study focused on two underexplored issues in research regarding (regressed) PE interest fit to contribute to vocational interest literature and optimize the use of PE interest fit in counseling settings. First, regressed PE interest fit is thus far applied within the educational context to predict academic achievement. The present study extends this application to the prediction of *study choice* in itself, which is established to be of comparable research significance in vocational interest literature (Rounds & Su, 2014; Stoll et al., 2017). Second, existing PE interest fit measures only acknowledge interest patterns that align with a specific environment (Allen & Robbins, 2010; Tracey et al., 2012). However, this rationale does not do conceptual justice to the reality of the study choice decision-making process, during which a student chooses *between* study programs that may differ more or less from each other (Feldman et al., 2001; Weidman, 2005). Choosing one study program implies not choosing another one. Hence, the interest patterns of the non-chosen study programs also matter. We therefore proposed logistic regressed PE interest fit (LRIF), a PE interest fit method that differentiates between interest patterns that *align* with a specific study program like Psychology, represented by students that choose the Psychology study program, and interest patterns that *divert* from this study program, represented by students that choose another study program.

The present study set out to address the differentiation of interest patterns in an open access study environment, where a student's suitable study choice can prove crucial to their study success (Nauta, 2010; Schelfhout et al., 2021). For this purpose, students benefit from receiving accurate and valid study advice, ideally developed in the most time-efficient manner possible. Data were therefore analyzed on $N = 14{,}175$ Flemish first-year university students, across the consecutive academic years 2016-2017, 2017-2018 and 2018-2019 and spread out over $n = 31$ independent study program subsets. Specifically, the present study investigated to what extent LRIF predicts study choice. Furthermore, we compared the predictive validity for study choice obtained through LRIF versus through Euclidean distance PE interest fit (EDF) (Wille et al., 2014) and correlation PE interest fit (CF) (Tracey et al., 2012). Unlike these latter two traditional continuous PE interest fit methods, LRIF does not require additional data collection to predetermine study program interest profiles and does account for the differentiation of interest patterns.

For the 31 study programs, the LRIF method obtains acceptable to outstanding model fits for the prediction of study choice. This finding implies that LRIF can effectively differentiate between students who choose a specific study program and those who choose another study program. The results are consistent with previous research that established positive associations between vocational interests and occupational/study choice on the one hand (Rounds & Su, 2014; Stoll et al., 2017), and between (regressed) PE interest fit and work- and study outcomes on the other hand (Nye, Butt, et al., 2018; Nye, Prasad, et al., 2018; Schelfhout, Bassleer, et al., 2022). Moreover, a very large mean LRIF difference is found for students who choose a particular study program compared with students who choose another study program. Thus, students that choose a specific study program better match with that environment, as indicated by their PE interest fit, in contrast to those who choose another study program. PE (interest) fit indeed builds on the premise that individuals seek a match with their (work or study) environment (De Cooman et al., 2009; Holland, 1997; Oh et al., 2018).

More importantly, LRIF rivals and even outperforms the more conventional continuous PE interest fit measures CF and EDF, respectively, in predicting study choice. The findings for LRIF are at least on par with those for CF, while the advantage over EDF is substantial. As such, these results correspond to our expectation of an at least equally predictive validity for study choice of LRIF versus CF and EDF, in which LRIF offers conceptual and practical advantages. Indeed, by also taking into account the differentiating interest patterns of a study program and making additional data collection superfluous (in contrast to the CF and EFD methods), LRIF succeeds in obtaining equivalent and even increased (compared to EDF) validity in capturing actual study choice. The present study's findings align with the viewpoint of researchers who endorse the utilization of regressed PE interest fit methods above congruence indices and traditional continuous PE interest fit methods (Edwards, 1993, 1994; Nye, Prasad, et al., 2018), with the understanding that both fit and misfit should be considered (De Cooman & Vleugels, 2022; Jansen & Kristof-Brown, 2006). Previous educational research also shows superior performance of regressed PE interest fit in predicting academic achievement when compared to congruence indices (Nye, Prasad, et al., 2018) and traditional continuous PE interest fit measures (Schelfhout, Bassleer, et al., 2022). However, these studies focus on academic achievement rather than study choice. Here, the fact that LRIF and CF outperform EDF in predictive validity for study choice is not surprising, just like CF also outperforms EDF in predicting career outcomes (Xu & Li, 2020). Note that the average CF of $r = .68$ for students with their study programs is also consistent with literature on open access study environments (Schelfhout et al., 2019). Considering that CF is one of the strongest predictors of study choice (Schelfhout et al., 2019; Su & Rounds, 2015), the results of the present study reveal that LRIF can match this performance and does not lag behind. The correlation of $r = .68$ (i.e., 46% shared variance) between LRIF and CF indeed indicate a strong association between these PE interest fit measures, while still maintaining a degree of distinctiveness. Importantly, in addition to these similar predictive capabilities of LRIF and CF for predicting study choice, LRIF also brings valuable theoretical and practical implications for the concept of PE interest fit.

Theoretical Implications

PE interest fit is traditionally conceptualized by comparing the vocational interests of students (or employees) with interest patterns that align with a specific study program (or job environment) (Allen & Robbins, 2010; Tracey et al., 2012). Examples of such conceptualizations are EDF (Wille et al., 2014) and CF (Tracey et al., 2012). To detect these interest patterns that align with specific environment like a study program, one often used method is the incumbent method, in which interest patterns of the students that choose the study program are representative of the study program as an environment (Allen & Robbins, 2010; Schneider, 1987). Typically, environment interest profiles established through this incumbent method averages out students' scores on the RIASEC dimensions, but do not consider the variance within and covariance between these interest dimensions (Edwards, 1993; van Vianen, 2018). For the present study, the results for EDF and CF are in line with literature regarding their descriptive statistics (Wille et al., 2014), their predictive power towards study choice (Schelfhout et al., 2019, 2021; Su & Rounds, 2015) and their intercorrelation (Tracey et al., 2012; Wille et al., 2014), further supporting the psychometric qualities of the present data set.

In contrast however, the present study and LRIF address study choice as a choice between study programs, which acknowledges the reality of the study choice process, especially in an open access study environment (OECD, 2022; Weidman, 2005). Indeed, the differentiation of study programs regarding interest patterns represents the study choice students make in selecting as well as rejecting study programs. By design, the LRIF method complements the incumbent philosophy that the people make the environment (De Cooman & Vleugels, 2022; Hoff et al., 2020). LRIF not only considers students that choose a study program as representatives of this specific study program, like EDF and CF, but also takes into account students that choose another study program as non-representatives of this specific study program. In such a way, PE interest fit differentiates between interest patterns that align with a specific environment and interest patterns that divert from this environment. To the best of our knowledge, such an interpretation of PE interest fit is previously unknown to literature.

Practical Implications

By introducing the differentiation of interest patterns as an additional element to PE interest fit, the present study also has practical implications. First, the presented LRIF method enables to calculate PE interest fit using a dichotomous variable like study choice as the criterion. A logistic regression is still linear in its parameters (Edwards, 1993; Nye, Prasad, et al., 2018), as illustrated by a logit transformation. Such an extension of the regressed PE interest fit methodology is not reported previously.

Second, of great value in time efficiency, the LRIF method removes the need for additional data collection to predetermine environment interest profiles when doing research on the PE interest fit of individuals with an environment. Indeed, to establish PE interest fit between a prospective student and a study program using a traditional (continuous) PE interest fit method, a measurement must first be conducted among a large number of already enrolled students, such as third-year students in that particular program. By contrast, the LRIF method can estimate environment interest profiles itself through the obtained regression coefficients, which is a feat that is inherent to all regressed PE interest fit methods in literature (Edwards, 1994; Schelfhout et al., 2022). Observing the results of the present study, the correlations between (1) the EDF/CF study program interest profiles estimated by averaging out the RIASEC dimension scores over third-year students in the study programs and (2) the LRIF study program interest profiles are strong, but not perfect (mean correlation of $r = .83$). This strong association further supports the LRIF method, which is more practically advantageous compared with other more conventional (continuous) PE interest fit methods, as a valid way to establish environment interest profiles for environments like a study program. However, the not perfect correlation observation indicates that the LRIF method also captures other information in addition to interest patterns that align with a specific environment. The most important difference with traditional PE interest fit methods is that LRIF also considers interest patterns that divert from a specific environment. We posit that this conceptual feature is responsible for the increased explained variance in study choice found through LRIF (especially relative to EDF).

Limitations and Future Research

Despite the compelling results towards choice prediction and profile generation, we acknowledge that the present study also has limitations that need to be addressed. One could have expected LRIF to have been a more accurate predictor of study choice compared to CF, considering that regressed PE interest fit methods impose fewer constraints on the data than traditional methods (Edwards, 1994; van Vianen, 2018). However, LRIF does offer a theoretical advantage by also incorporating what is outside a specific environment (De Cooman & Vleugels, 2022; Jansen & Kristof-Brown, 2006). In addition, the present study was conducted in a very specific (open access) higher education environment. Future research could also focus on the application of LRIF in predicting choices within work contexts. Furthermore, although the present study does feature 31 largely independent study program environments, the study does make use of only one dichotomous variable (i.e., study choice). Future studies using (L)RIF may investigate other important outcome variables in higher education and/or work environments. For these variables, distinctions can be made among attitudinal outcomes (e.g., satisfaction, commitment) (Oh et al., 2014), well-being (e.g., stress) (van Vianen, 2018), and behavioral outcomes (e.g., degree attainment and retention) (Van Iddekinge et al., 2011).

Conclusion

LRIF allows to accurately distinguish between students who choose a specific study program versus those who choose another study program, while rivaling and even outperforming more traditional PE interest fit measures like CF and EDF, respectively. The LRIF method evolves the PE interest fit concept from considering interest patterns that align with a specific environment towards differentiating between these interest patterns and those that divert from this specific environment. Moreover, LRIF does not require additional data collection to predetermine environment interest profiles, unlike traditional PE interest fit methods, yielding a time-efficient benefit in study and career counseling contexts.

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. https://doi.org/10.1037/a0017267

De Cooman, R., Gieter, S. De, Pepermans, R., Hermans, S., Bois, C. Du, Caers, R., & Jegers, M. (2009). Person–organization fit: Testing socialization and attraction–selection–attrition hypotheses. *Journal of Vocational Behavior*, *74*(1), 102–107. https://doi.org/10.1016/j.jvb.2008.10.010

De Cooman, R., & Vleugels, W. (2022). Person–Environment Fit: Theoretical Perspectives, Conceptualizations, and Outcomes. In *Oxford Research Encyclopedia of Business and Management*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190224851.013.377

Edwards, J. R. (1993). Problems with the Use of Profile Similarity Indices in the Study of Congruence in Organizational Research. *Personnel Psychology*, *46*(3), 641–665. https://doi.org/10.1111/j.1744-6570.1993.tb00889.x

Edwards, J. R. (1994). The Study of Congruence in Organizational Behavior Research: Critique and a Proposed Alternative. *Organizational Behavior and Human Decision Processes*, *58*(1), 51–100. https://doi.org/10.1006/obhd.1994.1029

Feldman, K. A., Ethington, C. A., & Smart, J. C. (2001). A Further Investigation of Major Field and Person-Environment Fit. *The Journal of Higher Education*, *72*(6), 670–698. https://doi.org/10.1080/00221546.2001.11777121

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.

Fonteyne, L., Duyck, W., & de Fruyt, F. (2017). Program-specific prediction of academic achievement on the basis of cognitive and non-cognitive factors. *Learning and Individual Differences*, *56*, 34–48. https://doi.org/10.1016/j.lindif.2017.05.003

Guan, Y., Deng, H., Fan, L., & Zhou, X. (2021). Theorizing person-environment fit in a changing career world: Interdisciplinary integration and future directions. *Journal of Vocational Behavior*, *126*, 103557. https://doi.org/10.1016/j.jvb.2021.103557

Hoff, K. A., Song, Q. C., Wee, C. J. M., Phan, W. M. J., & Rounds, J. (2020). Interest fit and job satisfaction: A systematic review and meta-analysis. *Journal of Vocational Behavior*, *123*, 103503. https://doi.org/10.1016/j.jvb.2020.103503

Holland, J. L. (1997). *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments (3rd ed.)*. Psychology Assessment Resources.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression*. John Wiley & Sons, Inc. https://doi.org/10.1002/0471722146

Jansen, K. J., & Kristof-Brown, A. (2006). Toward a Multidimensional Theory of Person-Environment Fit. *Journal of Managerial Issues*. https://www.jstor.org/stable/40604534

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a Unifying Social Cognitive Theory of Career and Academic Interest, Choice, and Performance. *Journal of Vocational Behavior*, *45*(1), 79–122. https://doi.org/10.1006/jvbe.1994.1027

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, *78*(3), 691–692. https://doi.org/10.1093/biomet/78.3.691

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. https://doi.org/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. https://doi.org/10.1016/j.jvb.2018.08.003

Nye, C. D., Perlus, J. G., & Rounds, J. (2018). Do ornithologists flock together? Examining the homogeneity of interests in occupations. *Journal of Vocational Behavior*, *107*, 195–208. https://doi.org/10.1016/j.jvb.2018.04.004

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. https://doi.org/10.1016/j.jvb.2017.10.012

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational Interests and Performance. *Perspectives on Psychological Science*, *7*(4), 384–403. https://doi.org/10.1177/1745691612449021

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2017). Interest congruence and performance: Revisiting recent meta-analytic findings. *Journal of Vocational Behavior*, *98*, 138–151. https://doi.org/10.1016/j.jvb.2016.11.002

OECD. (2017). *State of Higher Education 2015-16*. https://www.oecd.org/education/imhe/The%20State%20of%20Higher%20Education%20 2015-16.pdf

OECD. (2022). *Education at a Glance 2022*. OECD. https://doi.org/10.1787/3197152b-en

Oh, I.-S., Guay, R. P., Kim, K., Harold, C. M., Lee, J.-H., Heo, C.-G., & Shin, K.-H. (2014). Fit Happens Globally: A Meta-Analytic Comparison of the Relationships of Person-Environment Fit Dimensions with Work Attitudes and Performance Across East Asia, Europe, and North America. *Personnel Psychology*, *67*(1), 99–152. https://doi.org/10.1111/peps.12026

Oh, I.-S., Han, J. H., Holtz, B., Kim, Y. J., & Kim, S. (2018). Do birds of a feather flock, fly, and continue to fly together? The differential and cumulative effects of attraction, selection, and attrition on personality-based within-organization homogeneity and between-organization heterogeneity progression over ti. *Journal of Organizational Behavior*, *39*(10), 1347–1366. https://doi.org/10.1002/job.2304

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, *49*(12), 1373–1379. https://doi.org/10.1016/S0895-4356(96)00236-3

Rounds, J., Smith, T., & Hubert, L. (1999). *Development of Occupational Interest Profiles for O*NET*. https://www.onetcenter.org/reports/OIP.html

Rounds, J., & Su, R. (2014). The Nature and Power of Interests. *Current Directions in Psychological Science*, *23*(2), 98–103. https://doi.org/10.1177/0963721414522812

Sawilowsky, S. S. (2009). New Effect Size Rules of Thumb. *Journal of Modern Applied Statistical Methods*, *8*(2), 597–599. https://doi.org/10.22237/jmasm/1257035100

Schelfhout, S., Bassleer, M., Wille, B., Van Cauwenberghe, S., Dutry, M., Fonteyne, L., Dirix, N., Derous, E., De Fruyt, F., & Duyck, W. (2022). Regressed person-environment interest fit: Validating polynomial regression for a specific environment. *Journal of Vocational Behavior*, *136*, 103748. https://doi.org/10.1016/j.jvb.2022.103748

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

Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., Derous, E., De Fruyt, F., & Duyck, W. (2021). How interest fit relates to STEM study choice: Female students fit their choices better. *Journal of Vocational Behavior*, *129*, 103614. https://doi.org/10.1016/j.jvb.2021.103614

Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., Derous, E., De Fruyt, F., & Duyck, W. (2022). How accurately do program-specific basic skills predict study success in open access higher education? *International Journal of Educational Research*, *111*, 101907. https://doi.org/10.1016/j.ijer.2021.101907

Schneider, B. (1987). THE PEOPLE MAKE THE PLACE. *Personnel Psychology*, *40*(3), 437–453. https://doi.org/10.1111/j.1744-6570.1987.tb00609.x

Statistiek Vlaanderen. (2024, March 26). *Drop-out in het hoger onderwijs*. https://www.vlaanderen.be/statistiek-vlaanderen/onderwijs-en-vorming/drop-out-in-het-hoger-onderwijs

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. https://doi.org/10.1037/pspp0000117

Su, R., & Nye, C. D. (2017). *Interests and Person–Environment Fit* (Vol. 1). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199373222.003.0008

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*. https://doi.org/10.3389/fpsyg.2015.00189

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. https://doi.org/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. https://doi.org/10.1016/j.jvb.2011.03.005

Van Iddekinge, C. H., Roth, P. L., Putka, D. J., & Lanivich, S. E. (2011). Are you interested? A meta-analysis of relations between vocational interests and employee performance and turnover. *Journal of Applied Psychology*, *96*(6), 1167–1194. https://doi.org/10.1037/a0024343

van Vianen, A. E. M. (2018). Person–Environment Fit: A Review of Its Basic Tenets. *Annual Review of Organizational Psychology and Organizational Behavior*, *5*(1), 75–101. https://doi.org/10.1146/annurev-orgpsych-032117-104702

Weidman, J. C. (2005). Academic Disciplines: Holland's Theory and the Study of College Students and Faculty. *The Journal of Higher Education*, *76*(2), 232–234. https://doi.org/10.1080/00221546.2005.11778912

Wilkins, K. G., & Tracey, T. J. G. (2014). Person Environment Fit and Vocational Outcomes. In *Psycho-social Career Meta-capacities* (pp. 123–138). Springer International Publishing. https://doi.org/10.1007/978-3-319-00645-1_7

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. https://doi.org/10.1016/j.jvb.2013.12.001

Xu, H., & Li, H. (2020). Operationalize Interest Congruence: A Comparative Examination of Four Approaches. *Journal of Career Assessment*, *28*(4), 571–588. https://doi.org/10.1177/1069072720909825

6
Discussion

The shift from secondary to higher education presents a challenge for (prospective) students (Tett et al., 2017; van Daal et al., 2013). In Flanders, the higher education system features open access (see also Chapter 1, Footnote 1) (Eurydice, 2023; OECD, 2021a) and low enrollment fees (Kelchtermans & Verboven, 2010; OECD, 2021a). While this structure aims to promote broad access, students may encounter difficulties in the study choice decision-making process given the abundance of study options (Fonteyne, 2017, 2022). Consequently, higher education success rates can be affected (Fonteyne, 2017; Schelfhout, 2019), with costs for students, parents, and society (Fonteyne, 2017; OECD, 2022; Schelfhout, 2019).

In response, Ghent University developed SIMON (Study capacities and Interest MONitor), an online self-assessment tool to assist prospective students in making a suitable study choice that matches their interests and provide tailored guidance to first-year higher education students considering their competences (Fonteyne, 2017, 2022). This approach enables the timely identification and support of students who lack the basic skills needed for academic success (Fonteyne, 2017; Schelfhout, 2019), on which SIMON previous research shows promising results (e.g., Schelfhout, 2019).

Since its launch in 2012 and widespread use from 2015-2016, SIMON has collected data from over 70,000 students (Fonteyne, 2022). Continuous optimization of SIMON is recommended to ensure its effectiveness in meeting the (evolving) needs of and insights from research, education, and its actors. Indeed, research evolves and society changes, alongside the relatively stable academic achievement metrics (Statistiek Vlaanderen, 2024b, 2024a).

The present dissertation therefore leverages (recent) conceptual and methodological knowledge combined with SIMON data to refine the tool, aiming to improve students' academic trajectories, fill research gaps, and benefit educational practice. In doing so, Chapter 2 and Chapter 3 focused on the competence component (SIMON-C), while Chapter 4 covered the associated feedback module. In addition, Chapter 5 concentrated on the interest component (SIMON-I).

Competences

COVID-19 Pandemic

In Chapter 2, we examined COVID-19 pandemic's effects on academic achievement (i.e., % obtained ECTS credits), including potential learning losses among low SES students and interactions between the pandemic and other academic achievement determinants. Therefore, the four pre-pandemic cohorts 2015-2016 to 2018-2019, the one-third pandemic cohort 2019-2020, and the full pandemic cohort 2020-2021 were taken into account. Recognizing unexpected societal changes and understanding its impacts on academic achievement is indeed relevant for the continued implementation of the SIMON project. Also, the influence of the pandemic on academic achievement in higher education as well as interactions between the pandemic and student characteristics other than socio-economic status (SES) and sex (Betthäuser et al., 2023; Iterbeke & De Witte, 2021) are understudied in literature. Due to the comprehensive data from SIMON-C on student characteristics and academic achievement over pre- and pandemic years, we were able to gain valuable insights into the effects of a crisis (i.e., the COVID-19 pandemic) on higher education.

On average, students from the one-third pandemic cohort 2019-2020 show higher academic achievement compared with the pre-pandemic cohorts but with small effects. This finding contrasts with compulsory education, where severe learning losses are observed (Betthäuser et al., 2023). Modifications to learning materials, exams, tasks, and grading, possibly leading to increased study time for students, might have contributed to the enhanced academic achievement in higher education during this time (e.g., Bird et al., 2022; Rodríguez-Planas, 2022). Conversely, the full pandemic cohort 2020-2021 demonstrate the lowest academic achievement compared with the pre-pandemic cohorts 2015-2016 to 2017-2018.

However, the largest academic achievement difference is observed when contrasting this full pandemic cohort 2020-2021 to the one-third pandemic cohort 2019-2020, with cohort 2020-2021 achieving a small one-semester course less (i.e., three ECTS credits). Factors contributing to this decline in 2020-2021 versus 2019-2020 may include these students' complete reliance on (partial) distance learning during their first year of higher education and during the completion of their last year of secondary education. As a result, the full pandemic cohort 2020-2021 also experienced a lack of social integration in the academic context (Kassarnig et al., 2018; Rayle & Chung, 2007), unlike the one-third pandemic cohort 2019-2020 who still had exposure to some form of higher education functioning. Furthermore, only SES seems to moderate the relationship between the pandemic and academic achievement. The largest academic achievement gap between low SES and high SES students is found for the full pandemic cohort 2020-2021 (where also the largest variance occurs), with a difference of achieving a larger one-semester course (i.e., six ECTS credits). Especially during the full pandemic year, accessible and affordable e-learning and a supportive environment were important (Azevedo et al., 2022; Hammerstein et al., 2021; OECD, 2021). However, the availability of these essential resources is more challenging among low SES students (Azevedo et al., 2022; OECD, 2021), who also tend to face higher COVID-19 infection and mental distress risks (Anderson et al., 2020). In sum, the results imply that the observed learning losses in higher education are rather small, but acknowledgement for the vulnerability of and support for low SES students should not be overlooked.

Dutch Language Proficiency

In Chapter 3, we introduced the Ghent University Language Screening (GULS) as a fully open access Dutch post-entry language assessment (PELA). We evaluated GULS's construct validity, reliability and predictive validity for academic achievement (i.e., GPA and study success) across 16 bachelor's study programs using data from the three-year periods

2017-2018 to 2019-2020 and 2020-2021 to 2022-2023, as well as its program-specific predictive validity for academic achievement across these six academic years. A similar study has been conducted for the basic mathematics test in SIMON, but not yet for GULS which specifically assesses academic language proficiency in terms of reading comprehension of first-year higher education students. Further, non-English PELAs are scarce, and frequently lack professional construct validation (Knoch & Elder, 2013) and analyses of the extent to which language proficiency differently impacts academic achievement in distinct fields of study (Elder, 2017; Read, 2016). To the best of our knowledge, there are also no publicly available Dutch PELAs at present. SIMON-C allowed the assessment of individual factors, such as language proficiency, on academic achievement across several academic years and study programs within an ecologically valid context.

GULS with 18 items demonstrates good construct validity and reliability. The test information is at its best for students with lower language proficiency scores, aligning with the purpose of PELAs to identify and offer language support to at-risk students (Elder, 2017; Knoch & Elder, 2013; Read, 2016). Additionally, GULS shows modest predictive validity for academic achievement across study programs, consistent with previous research on language assessment for admission in English-medium higher education institutions (for meta-analyses, see e.g., Ihlenfeldt & Rios, 2023; Clinton-Lisell et al., 2022) and (non-)Dutch PELAs (e.g., De Wachter et al., 2013; Heeren et al., 2021; van Dijk, 2015). Moreover, GULS predicts academic achievement above and beyond background and other cognitive student characteristics. Similarly, the (incremental) predictive validity of GULS for academic achievement is found both across basic mathematics as well as advanced mathematics study programs. Notably, we also confirm the differential contribution of GULS to academic achievement at the program-specific level. The study program Applied Language Studies is distinctive, while GULS plays no role in predicting academic achievement for Biochemistry and Biotechnology, and Engineering Technology. Our findings indicate that, on the one hand, GULS can be used for conducting population-level research in higher education across and within diverse study programs. On the other hand, GULS can provide non-binding Dutch language proficiency advice for first-year higher education students, thereby helping identify and support at-risk students. GULS can also offer tailored advice for specific advanced mathematics study programs, with caution for basic ones (Evers et al., 2009; Nunnally & Bernstein, 1994).

Feedback

Student Feedback Engagement

In Chapter 4, we used a model based on the Theory of Planned Behavior (TPB) to investigate student feedback engagement (i.e., intentional and behavioral feedback engagement) among first-year university students who receive feedback on their first-year predicted chance of study success (i.e., (very) low, (fairly) high), supplemented with recommendations for remediation/competence training activities. The inclusion of feedback self-efficacy allowed for a better understanding of determinants and underlying mechanisms of student feedback engagement. As such, SIMON as well as the broader educational field can gain insights into designing interventions that promote student feedback engagement. Additionally, literature calls for more studies in this area (Panadero & Lipnevich, 2022; Winstone & Nash, 2023). The feedback module associated with SIMON-C enabled us to monitor first-year higher education students over time, evaluate engagement with feedback that is not experimentally and/or virtually manipulated, and determine how to improve student feedback engagement.

All findings, except one, are consistent with the principles of the TPB (Ajzen, 1991, 2012). Students' intention to engage with the received feedback positively influences their actual feedback engagement, which aligns with previous meta-analyses across diverse disciplines (e.g., Riebl et al., 2015). The readiness to dedicate effort and time in performing a behavior is reflected in an individual's intentions (Ajzen, 1991; Handley et al., 2011). Additionally, feedback self-efficacy has a positive effect on intentional feedback engagement, but only indirectly through the intentions on behavioral feedback engagement. Thus, for students' belief in their own feedback engagement capacities to impact the behavior of feedback engagement, their intentions towards this behavior must be present. While the TPB typically suggests partial mediation in this regard, this assumption may change based on the specific context (Ajzen, 1991, 2012). Further, students who received positive feedback in terms of a (fairly) high versus a (very) low chance of study success in their enrolled study program show higher feedback self-efficacy, corresponding with literature (Brown et al., 2016; Peifer et al., 2020). The emotional/psychological state that can be boosted by success feedback is indeed a recognized source of self-efficacy (Bandura, 1977). Moreover, feedback self-efficacy fully mediates and thus plays an essential role in the relationship between the received feedback and intentional feedback engagement. Students with higher confidence in their ability to engage with feedback tend to regard negative/failure feedback more as valuable learning opportunities (Adams et al., 2020; Putwain et al., 2013). Hence, we suggest that educational interventions could focus on the enhancement of feedback self-efficacy to improve student feedback engagement.

Interests

Logistic Regressed Person-Environment Interest Fit

Finally, in Chapter 5, we investigated to what extent logistic regressed person-environment (PE) interest fit (LRIF) predicts study choice, including 31 study programs. In addition, we compared the predictive validity for study choice obtained through LRIF versus through Euclidean distance PE interest fit (EDF) and correlation PE interest fit (CF). In fact, PE interest fit measurement methods that eliminate the need for extra data collection to obtain environment interest profiles (Edwards, 1994; Schelfhout et al., 2022) like LRIF, and hold potential for providing more accurate study advice, are valuable in counseling contexts such as SIMON. Furthermore, the proposed LRIF method focuses on study choice (Rounds & Su, 2014; Stoll et al., 2017) and considers both the interest patterns that align with and those that divert from a specific environment (Feldman et al., 2001; Weidman, 2005), two underexplored issues in current (regressed) PE interest fit research. SIMON-I data from various academic years and study programs made it possible to address more methodological issues as well, leading to the introduction and investigation of the practically advantageous LRIF method.

LRIF is found to effectively differentiate between students who choose a specific study program and those who choose another study program, aligning with previous studies on vocational interests and environment choice (Rounds & Su, 2014; Stoll et al., 2017), and the predictive validity of (regressed) PE interest fit for work- and study outcomes (Nye, Butt, et al., 2018; Nye, Prasad, et al., 2018; Schelfhout et al., 2022). Moreover, LRIF surpasses EDF in predicting study choice and equals CF, which is a strong predictor of study choice (Schelfhout et al., 2019; Su & Rounds, 2015) that also exceeds EDF in predicting career outcomes (Xu & Li, 2020). LRIF can therefore function as a full-fledged alternative for traditional PE interest fit measures in study (and career) counseling settings, providing the benefits of not requiring additional data collection to predetermine environment interest profiles and considering the differentiation of interest patterns.

Future Guidelines

First, the present dissertation relied on data from a relative open access higher education institution, Ghent University. While we made efforts to include various study programs and faculties from this university, the findings' generalizability to other educational institutions is still unclear. However, in Chapter 3, we see that the contribution of language proficiency to academic achievement across study programs falls within the same range as the contributions found in other studies that use both open (e.g., De Wachter et al., 2013; Heeren et al., 2021; van Dijk, 2015) and closed access systems (Ihlenfeldt & Rios, 2023; Clinton-Lisell et al., 2022). Nevertheless, a comparison with other open access and/or closed access higher education systems following the same methodology is relevant to accurately identify potential differences regarding various educational topics.

Second, our approach to assessing student characteristics was based on self-reported data. Although we thoroughly considered a range of student attributes where relevant and feasible, self-reporting introduces the possibility of social desirability bias, particularly regarding non-cognitive factors. This issue is notably relevant in Chapter 4, where constructs like feedback self-efficacy and both intentional and behavioral feedback engagement were measured via self-reports. For behavioral feedback engagement, in particular, an (additional) objective measure would have been more advantageous, providing a clearer picture of actual behavior (Ajzen, 1985, 2020). Furthermore, SIMON refers the Monitoring Service and Study Advice Department of the university, where students can seek further support and guidance based on their SIMON feedback. The integration of data from these services into SIMON research and practice could offer significant value. For example, knowing whether and why a student attended a session provided by the Monitoring Service and/or Study Advice Department would enhance our ability to assess student feedback engagement and its underlying mechanisms.

Third, in terms of academic achievement measures, Chapter 2 addressed study success, while Chapter 3 included GPA as an additional metric. The use of GPA is debated in literature (Ihlenfeldt & Rios, 2023), and timely graduation seems to be gaining traction as a measure of academic achievement (Kim, 2023; Moraga-Pumarino et al., 2023). Timely graduation can be defined as completing a bachelor's degree within the standard duration of three years or with a one-year extension (OECD, 2023). Although SIMON has recently acquired data on students who completed their bachelor's degree in three or four years and participated in SIMON during their first year, this data was not sufficiently available for inclusion in the present dissertation. Future SIMON research could incorporate timely graduation as a key metric to provide a more comprehensive assessment of academic achievement.

Fourth, we used data from the ‘SIMON Zegt’ application and thus from first-year higher education students (and vocational interest data of third-year bachelor’s students in Chapter 5). This data also informs study advice for last-year secondary education students who engage with the ‘Vraag het aan SIMON’ application. However, through this latter application, data on these last-year secondary students are also available but have yet to be explored. Future research could greatly benefit from following these students throughout their higher education journey. Indeed, such longitudinal studies could yield valuable insights into the educational paths ultimately chosen by these students after their participation in ‘Vraag het aan SIMON’, the contributing factors to their study choice decisions, and the subsequent impact on their academic achievement.

Finally, we already acknowledged the importance of considering societal changes and aligning the SIMON project as necessary to provide up-to-date study advice. This recognition is exemplified by our focus on the COVID-19 pandemic in Chapter 2. Meanwhile, our current society is increasingly influenced by the growing presence of artificial intelligence (AI), which is inevitably impacting learning and teaching (García-Martínez et al., 2023; Shahzad et al., 2024). In the future, awareness and monitoring the integration of AI in education will therefore be essential. Thinking innovatively, AI could even be integrated into SIMON applications, enabling students to ask specific questions directly to an AI system regarding their received SIMON feedback. Additionally, optional (e.g., for ‘Economics’ and ‘Biology’) and mandatory positioning tests (e.g., for ‘Pharmaceutical Sciences’ and ‘Industrial Engineering’) are recently introduced for the majority of academic study programs. Both tests are non-binding and thus continue to provide access to higher education (unlike admission exams). However, if students do not pass a mandatory positioning test, they are required to undertake a remedial program (Universiteit Gent, 2024). Similarly to the SIMON project (especially regarding the 'SIMON zegt' application), these tests thus play a role in students’ study (re)orientation and remediation processes in/towards higher education, an issue recommended for future monitoring.

Conclusion

In sum, we could say that the present dissertation provides ‘guidance for better study guidance’ towards/in higher education. Indeed, we focused on various aspects of SIMON to contribute to its optimization for the improvement of students’ academic trajectories, thereby providing insights for research and educational practice as well. First, we showed that pandemic-induced learning losses in higher education are relatively minor and less pronounced than in compulsory education. Slight SES-related differences occur, highlighting the need for continued recognition of and support for low SES students. Second, the open access Dutch language proficiency test in terms of reading comprehension, GULS, is a valid and reliable tool for future population-level research and study (re)orientation and remediation advice, with differential contributions to academic achievement depending on the study program. Third, we find that educational interventions can concentrate on improving students’ feedback self-efficacy to enhance their engagement with received (SIMON) feedback. Last, the introduced LRIF method provides an alternative approach to assess PE interest fit, predicting study choice at least as accurate as traditional methods. However, LRIF refines the PE interest fit concept by considering the differentiation of interest patterns, and not requiring additional data collection to obtain environment interest profiles. Hence, this method offers time-efficient benefits for study (and career) counseling settings like SIMON.

References

Abunawas, M. (2014). *A meta-analytic investigation of the predictive validity of the Test of English as a Foreign Language (TOEFL) scores on GPA*. https://oaktrust.library.tamu.edu/bitstream/handle/1969.1/154156/ABUNAWAS-DISSERTATION-2014.pdf?sequence=1

Adams, A.-M., Wilson, H., Money, J., Palmer-Conn, S., & Fearn, J. (2020). Student engagement with feedback and attainment: the role of academic self-efficacy. *Assessment & Evaluation in Higher Education*, *45*(2), 317–329. https://doi.org/10.1080/02602938.2019.1640184

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckman (Eds.), *Action-control: From cognition to behavior* (pp. 11–39). Springer.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Ajzen, I. (2012). The Theory of Planned Behavior. In *Handbook of Theories of Social Psychology: Volume 1* (pp. 438–459). SAGE Publications Ltd. https://doi.org/10.4135/9781446249215.n22

Ajzen, I. (2020). The theory of planned behavior: Frequently asked questions. *Human Behavior and Emerging Technologies*, *2*(4), 314–324. https://doi.org/10.1002/hbe2.195

Anderson, G., Frank, J. W., Naylor, C. D., Wodchis, W., & Feng, P. (2020). Using socioeconomics to counter health disparities arising from the covid-19 pandemic. *BMJ*, m2149. https://doi.org/10.1136/bmj.m2149

Azevedo, P., Gutierrez, M., de Hoyos, R., & Saavedra, J. (2022). The Unequal Impacts of COVID-19 on Student Learning. In F. M. Reimers (Ed.), *Primary and Secondary*

Education During Covid-19. Disruptions to Educational Opportunity During a Pandemic (pp. 421–459). https://doi.org/10.1007/978-3-030-81500-4_16

Bandura, A. (1977). Self-efficacy: Toward a Unifying Theory of Behavioral Change. In *Psychological Review* (Vol. 84, Issue 2). https://search.proquest.com/docview/614275783?accountid=11077

Betthäuser, B. A., Bach-Mortensen, A. M., & Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. *Nature Human Behaviour*, *7*(3), 375–385. https://doi.org/10.1038/s41562-022-01506-4

Bird, K. A., Castleman, B. L., & Lohner, G. (2022). Negative Impacts From the Shift to Online Learning During the COVID-19 Crisis: Evidence From a Statewide Community College System. *AERA Open*, *8*, 233285842210812. https://doi.org/10.1177/23328584221081220

Brown, G. T. L., Peterson, E. R., & Yao, E. S. (2016). Student conceptions of feedback: Impact on self-regulation, self-efficacy, and academic achievement. *British Journal of Educational Psychology*, *86*(4), 606–629. https://doi.org/10.1111/bjep.12126

De Wachter, L., Heeren, J., Marx, S., & Huyghe, S. (2013). "Taal: een noodzakelijke, maar niet de enige voorwaarde tot studiesucces: De correlatie tussen de resultaten van een taalvaardigheidstoets en de slaagcijfers bij eerstejaarsstudenten aan de KU Leuven." [Language Proficiency: A Necessary, but Not the Only, Condition for Study Success: A Correlation between the Results of a Language Proficiency Test and Academic Achievement of First-Year Students.]. *Levende Talen Tijdschrift*, *14*(4), 28–36. https://lt-tijdschriften.nl/ojs/index.php/ltt/article/view/549

Edwards, J. R. (1994). The Study of Congruence in Organizational Behavior Research: Critique and a Proposed Alternative. *Organizational Behavior and Human Decision Processes*, *58*(1), 51–100. https://doi.org/10.1006/obhd.1994.1029

Elder, C. (2017). Language Assessment in Higher Education. In *Language Testing and Assessment* (pp. 271–286). Springer International Publishing. https://doi.org/10.1007/978-3-319-02261-1_35

Eurydice. (2023, November 27). *Belgium - Flemish Community*. https://eurydice.eacea.ec.europa.eu/national-education-systems/belgium-flemish-community/bachelor

Feldman, K. A., Ethington, C. A., & Smart, J. C. (2001). A Further Investigation of Major Field and Person-Environment Fit. *The Journal of Higher Education*, *72*(6), 670–698. https://doi.org/10.1080/00221546.2001.11777121

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.

Fonteyne, L. (2022). SIMON biedt handvatten voor studiekeuze en -succes. *TH&MA (DEN HAAG)*, *2022*(4), 14–18. http://hdl.handle.net/1854/LU-01GXAP2CGKG5SKH968VKDG4F3A

Gagen, T. (2019). *The predictive validity of IELTS scores: A meta-analysis*. https://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=8762&context=etd

García-Martínez, I., Fernández-Batanero, J. M., Fernández-Cerero, J., & León, S. P. (2023). Analysing the Impact of Artificial Intelligence and Computational Sciences on Student Performance: Systematic Review and Meta-analysis. *Journal of New Approaches in Educational Research*, *12*(1), 171. https://doi.org/10.7821/naer.2023.1.1240

Gonzalez, T., de la Rubia, M. A., Hincz, K. P., Comas-Lopez, M., Subirats, L., Fort, S., & Sacha, G. M. (2020). Influence of COVID-19 confinement on students' performance in higher education. *PLOS ONE*, *15*(10), e0239490. https://doi.org/10.1371/journal.pone.0239490

Hammerstein, S., König, C., Dreisörner, T., & Frey, A. (2021). Effects of COVID-19-Related School Closures on Student Achievement-A Systematic Review. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.746289

Handley, K., Price, M., & Millar, J. (2011). Beyond 'doing time': investigating the concept of student engagement with feedback. *Oxford Review of Education*, *37*(4), 543–560. https://doi.org/10.1080/03054985.2011.604951

Hauspie, C., Schelfhout, S., Dirix, N., Fonteyne, L., Janse, M., Szmalec, A., Vereeck, A., & Duyck, W. (2024). Does Studying Latin in Secondary Education Predict Study Achievement in Academic Higher Education? *Language Learning*. https://doi.org/10.1111/lang.12639

Heeren, J., Speelman, D., & De Wachter, L. (2021). A practical academic reading and vocabulary screening test as a predictor of achievement in first-year university students: implications for test purpose and use. *International Journal of Bilingual Education and Bilingualism*, *24*(10), 1458–1473. https://doi.org/10.1080/13670050.2019.1709411

Iglesias-Pradas, S., Hernández-García, Á., Chaparro-Peláez, J., & Prieto, J. L. (2021). Emergency remote teaching and students' academic performance in higher education during the COVID-19 pandemic: A case study. *Computers in Human Behavior*, *119*, 106713. https://doi.org/10.1016/j.chb.2021.106713

Ihlenfeldt, S. D., & Rios, J. A. (2023). A meta-analysis on the predictive validity of English language proficiency assessments for college admissions. *Language Testing*, *40*(2), 276–299. https://doi.org/10.1177/02655322221112364

Kassarnig, V., Mones, E., Bjerre-Nielsen, A., Sapiezynski, P., Dreyer Lassen, D., & Lehmann, S. (2018). Academic performance and behavioral patterns. *EPJ Data Science*, *7*(1), 10. https://doi.org/10.1140/epjds/s13688-018-0138-8

Kelchtermans, S., & Verboven, F. (2010). Participation and study decisions in a public system of higher education. *Journal of Applied Econometrics*, *25*(3), 355–391. https://doi.org/10.1002/jae.1087

Kim, J. (2023). Unveiling Barriers to Timely Graduation and Strategies for Enhancing College Student Academic Completion. In *Research Highlights in Language, Literature and Education Vol. 7* (pp. 160–169). B P International (a part of SCIENCEDOMAIN International). https://doi.org/10.9734/bpi/rhlle/v7/5702C

Knoch, U., & Elder, C. (2013). A framework for validating post-entry language assessments (PELAs). *Papers in Language Testing and Assessment*, *2*(2), 48–66. https://arts.unimelb.edu.au/__data/assets/pdf_file/0007/1771306/4_Knoch_Elder_0.pdf

Moraga-Pumarino, A., Salvo-Garrido, S., & Polanco-Levicán, K. (2023). Profiles of University Students Who Graduate on Time: A Cohort Study from the Chilean Context. *Behavioral Sciences*, *13*(7), 582. https://doi.org/10.3390/bs13070582

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. https://doi.org/10.1016/j.jvb.2018.08.003

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. https://doi.org/10.1016/j.jvb.2017.10.012

OECD. (2021a). *Resourcing Higher Education in the Flemish Community of Belgium, Higher Education*. https://doi.org/https://doi.org/10.1787/3f0248ad-en

OECD. (2021b). *The State of Global Education - 18 Months into the Pandemic*. https://doi.org/10.1787/1a23bb23-en

OECD. (2022). *Education at a Glance 2022*. OECD. https://doi.org/10.1787/3197152b-en

OECD. (2023). *Education at a Glance 2023*. OECD. https://doi.org/10.1787/e13bef63-en

Panadero, E., & Lipnevich, A. A. (2022). A review of feedback models and typologies: Towards an integrative model of feedback elements. *Educational Research Review*, *35*, 100416. https://doi.org/10.1016/j.edurev.2021.100416

Peifer, C., Schönfeld, P., Wolters, G., Aust, F., & Margraf, J. (2020). Well Done! Effects of Positive Feedback on Perceived Self-Efficacy, Flow and Performance in a Mental Arithmetic Task. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01008

Putwain, D., Sander, P., & Larkin, D. (2013). Academic self-efficacy in study-related skills and behaviours: Relations with learning-related emotions and academic success. *British Journal of Educational Psychology*, *83*(4), 633–650. https://doi.org/10.1111/j.2044-8279.2012.02084.x

Rayle, A. D., & Chung, K.-Y. (2007). Revisiting First-Year College Students' Mattering: Social Support, Academic Stress, and the Mattering Experience. *Journal of College Student Retention: Research, Theory & Practice*, *9*(1), 21–37. https://doi.org/10.2190/X126-5606-4G36-8132

Read, J. (2016). *Some Key Issues in Post-Admission Language Assessment* (pp. 3–20). https://doi.org/10.1007/978-3-319-39192-2_1

Riebl, S. K., Estabrooks, P. A., Dunsmore, J. C., Savla, J., Frisard, M. I., Dietrich, A. M., Peng, Y., Zhang, X., & Davy, B. M. (2015). A systematic literature review and meta-analysis: The Theory of Planned Behavior's application to understand and predict nutrition-related behaviors in youth. *Eating Behaviors*, *18*, 160–178. https://doi.org/10.1016/j.eatbeh.2015.05.016

Rodríguez-Planas, N. (2022). COVID-19, college academic performance, and the flexible grading policy: A longitudinal analysis. *Journal of Public Economics*, *207*, 104606. https://doi.org/10.1016/j.jpubeco.2022.104606

Rounds, J., & Su, R. (2014). The Nature and Power of Interests. *Current Directions in Psychological Science*, *23*(2), 98–103. https://doi.org/10.1177/0963721414522812

Schelfhout, S. (2019). *The Influence of Study Interests and (Non-)Cognitive Predictors on Study Choice and Study Achievement in Flemish Higher Education* [Dissertation]. Ghent University.

Schelfhout, S., Bassleer, M., Wille, B., Van Cauwenberghe, S., Dutry, M., Fonteyne, L., Dirix, N., Derous, E., De Fruyt, F., & Duyck, W. (2022). Regressed person-environment interest fit: Validating polynomial regression for a specific environment. *Journal of Vocational Behavior*, *136*, 103748. https://doi.org/10.1016/j.jvb.2022.103748

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

Shahzad, M. F., Xu, S., Lim, W. M., Yang, X., & Khan, Q. R. (2024). Artificial intelligence and social media on academic performance and mental well-being: Student perceptions of positive impact in the age of smart learning. *Heliyon*, *10*(8), e29523. https://doi.org/10.1016/j.heliyon.2024.e29523

Statistiek Vlaanderen. (2024a, March 26). *Studieduur in het hoger onderwijs*. https://www.vlaanderen.be/statistiek-vlaanderen/onderwijs-en-vorming/studieduur-in-het-hoger-onderwijs

Statistiek Vlaanderen. (2024b, March 26). *Studierendement in het hoger onderwijs*. https://www.vlaanderen.be/statistiek-vlaanderen/onderwijs-en-vorming/studierendement-in-het-hoger-onderwijs

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. https://doi.org/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*. https://doi.org/10.3389/fpsyg.2015.00189

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. https://doi.org/10.1007/s10734-016-0101-1

Universiteit Gent. (2024). *Positioning tests and admission exams*. https://www.ugent.be/prospect/en/administration/application/requirement/admission-bachelor/positioningtests-admissionexams.htm

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

van Dijk, T. (2015). Tried and tested. *Tijdschrift Voor Taalbeheersing*, *37*(2), 159–186. https://doi.org/10.5117/TVT2015.2.VAND

Weidman, J. C. (2005). Academic Disciplines: Holland's Theory and the Study of College Students and Faculty. *The Journal of Higher Education*, *76*(2), 232–234. https://doi.org/10.1080/00221546.2005.11778912

Winstone, N. E., & Nash, R. A. (2023). Toward a cohesive psychological science of effective feedback. *Educational Psychologist*, *58*(3), 111–129. https://doi.org/10.1080/00461520.2023.2224444

Wongtrirat, R. (2010). *English language proficiency and academic achievement of international students: A meta-analysis*. https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1183&context=efl_etds

Xu, H., & Li, H. (2020). Operationalize Interest Congruence: A Comparative Examination of Four Approaches. *Journal of Career Assessment*, *28*(4), 571–588. https://doi.org/10.1177/1069072720909825

2A
Higher Education Organization During the COVID-19 Pandemic

Figure A1

Pandemic Matrix – Flemish Universities – July 2020

Activity	Code Green	Code Yellow	Code Orange	Code Red
Education in auditoria	No restrictions	Occupancy rate 1 to 2 with mouth mask obligation *or* Occupancy rate 1 to 5 without mouth mask obligation	Occupancy rate 1 to 5 with mouth mask obligation	Not possible/replaced by distance education
Education in small groups	No restrictions	Occupancy rate 1 to 2 with mouth mask obligation *or* Occupancy rate 1 to 5 without mouth mask obligation	Occupancy rate 1 to 2 with mouth mask obligation *or* Occupancy rate 1 to 5 without mouth mask obligation	Not possible/replaced by distance education
Practica and lab excercises	No restrictions	Occupancy rate 1 to 1 with mouth mask obligation	Occupancy rate 1 to 1 with mouth mask obligation	Occupancy rate 1 to 1 with mouth mask obligation
Internships	No restrictions	Safety regulations of the internship place apply	Safety regulations of the internship place apply	Safety regulations of the internship place apply
Incoming student mobility	No restrictions	Subject to federally and internationally established travel policies	Subject to federally and internationally established travel policies	Subject to federally and internationally established travel policies
Outgoing student mobility	No restrictions	Subject to federally and internationally established travel policies	Subject to federally and internationally established travel policies	Subject to federally and internationally established travel policies

Additional safety regulations (e.g., regarding hand hygiene, disinfection of rooms, ventilation...) are established in institution-specific safety protocols on the basis of a local risk analysis.

Note. Adopted from (VLIR, 2020) and translated to English.

Figure A2

Timetable Organization Teaching/Learning Activities Academic Year 2019-2020 (One Third Pandemic Year)

Pandemic Levels	’19-‘20													
	Mar.		Apr.			May		Jun.	Jul.		Aug.		Sep.	
	16/03 – 31/03	1/04 – 5/04	6/04 – 19/04	20/04 – 30/04	1/05 – 17/05	18/05 – 24/05	25/05 – 31/05	1/06 – 30/06	1/07 – 5/07	6/07 – 31/07	1/08 – 16/08	17/08 – 31/08	1/09 – 13/09	14/09 – 20/09
Code green														
Code yellow		Easter Holidays					First-chance exam period – SEM 2			Summer Holidays		Second-chance exam period		Free
Code orange														
Code red														

Note. SEM 2 = second semester. Closing days due to statutory public holidays and bridging days are included, but not mentioned explicitly. The university is always closed on Sundays. Each semester consists of twelve weeks of education activities followed by one week of catch-up activities.

Figure A3

Timetable Organization Teaching/Learning Activities First Semester Academic Year 2020-2021 (Full Pandemic Year)

Pandemic Levels	’20-‘21							
	Sep.	Oct.		Nov.	Dec.			Jan.
	21/09 – 30/09	1/10 – 25/10	26/10 – 31/10	1/11 – 30/11	1/12 – 20/12	21/12 – 31/12	1/01 – 3/01	4/01 – 31/01
Code green								
Code yellow						Christmas Holidays		First-chance exam period – SEM 1
Code orange								
Code red								

Note. SEM 1 = first semester. Closing days due to statutory public holidays and bridging days are included, but not mentioned explicitly. The university is always closed on Sundays. Each semester consists of twelve weeks of education activities followed by one week of catch-up activities.

Figure A4

Timetable Organization Teaching/Learning Activities Second Semester Academic Year 2020-2021 (Full Pandemic Year)

Pandemic Levels	’20-‘21														
	Feb.		Mar.		Apr.		May		Jun.	Jul.		Aug.		Sep.	
	1/02 – 7/02	8/02 – 28/02	1/03 – 31/03	1/04 – 4/04	5/04 – 18/04	19/04 – 30/04	1/05 – 23/05	24/05 – 31/05	1/06 – 30/06	1/07 – 4/07	5/07 – 31/07	1/08 – 15/08	16/08 – 31/08	1/09 – 12/09	13/09 – 26/09
Code green															
Code yellow	Free				Easter Holidays			First-chance exam period – SEM 2			Summer Holidays		Second-chance exam period		Free
Code orange															
Code red															

Note. SEM 2 = second semester. Closing days due to statutory public holidays and bridging days are included, but not mentioned explicitly. The university is always closed on Sundays. Each semester consists of twelve weeks of education activities followed by one week of catch-up activities.

References

VLIR. Pandemiematrix Vlaamse Universiteiten (2020, July 10). VLIR. https://vlir.be/nieuws/pandemiematrix-vlaamse-universiteiten/

2B

Extended Results

Table B1

Sample Sizes Students Registered for the First Time in Higher Education per Cohort

Cohort	*n*	Sex		Socio-Economic Status		Education Type Secondary Education	
		Male	Female	High SES	Low SES	General	Technical
2015-2016	3,631	1,640	1,991	2,783	848	3,396	235
2016-2017	3,980	1,732	2,248	3,120	860	3,717	263
2017-2018	3,851	1,505	2,346	3,075	776	3,610	241
2018-2019	4,330	1,804	2,526	3,404	9268	4,044	286
2019-2020	3,583	1,469	2,114	2,684	899	3,395	188
2020-2021	5,029	2,106	2,923	3,689	1,340	4,671	358
Total	24,404	10,256	14,148	18,755	5,649	22,833	1,571

Table B2

Sample Sizes Students Registered for the First Time in Higher Education per Study Program and per Cohort

Study Program	*n*					
	2015-2016	2016-2017	2017-2018	2018-2019	2019-2020	2020-2021
Psychology	271	365	442	400	299	468
Communication Sciences	75	108	95	88	63	69
Mathematics	20	24	22	20	17	23
Educational Sciences	62	96	109	120	97	87
Political Sciences	/	73	70	61	57	77
Law	340	265	336	304	215	394
Sociology	29	30	33	41	37	62
Criminological Sciences	92	127	162	156	122	249
Speech Language and Hearing Sciences	48	54	56	67	40	45
Physical Education and Movement Sciences	25	35	44	58	56	65
Philosophy	6	6	15	16	7	25
Linguistics and Literature	126	131	146	129	104	107
East European Languages and Cultures	6	5	11	13	5	7
History	58	67	58	70	45	79
Oriental Languages and Cultures	20	26	45	35	33	43
Moral Sciences	3	3	9	8	3	8
Art History	17	24	19	28	21	25
Archeology	14	12	18	18	18	26
African Studies	9	4	4	6	4	5
Veterinary Medicine	93	112	128	/	104	150
Physical Therapy and Motor Rehabilitation	294	312	253	277	292	360
Pharmaceutical Sciences	154	191	182	214	132	238
Bioscience Engineering	173	181	216	197	192	224
Bioscience Engineering Technology	4	17	3	1	/	/
Economics, Applied Economics and Business Economics	357	324	239	410	300	387

Study Program	*n*					
	2015-2016	2016-2017	2017-2018	2018-2019	2019-2020	2020-2021
Business Administration	236	282	183	320	260	382
Biomedical Sciences	122	100	113	197	116	224
Engineering - Architecture	63	56	42	61	76	109
Engineering	215	203	172	281	274	264
Bio Sciences	53	52	47	61	/	61
Engineering Technology	311	306	253	273	242	291
Applied Language Studies	111	113	119	113	80	115
Biochemistry and Biotechnology	48	63	57	62	69	80
Biology	30	41	32	34	56	65
Chemistry	26	32	24	28	26	38
Physics and Astronomy	24	41	25	34	32	70
Geology	9	13	6	24	12	9
Geography and Geomatics	14	11	7	15	8	16
Computer Sciences	27	33	33	40	30	41
Public Administration and Management	46	42	23	44	39	41
Total	3,631	3,980	3,851	4,330	3,583	5,029

Table B3

Descriptives and Multilevel Analysis Results with Cohort as Fixed Factor and Higher Education Study Program as Random Factor

	n	*M*	*SE*	95% Confidence Interval		*F*	R_m^2
				Lower Bound	Upper Bound		
Fixed Effects							
Cohort						11.04***	.002
2015-2016	3,631	68.95	1.41	66.13	71.78		
2016-2017	3,980	69.32	1.39	66.52	72.12		
2017-2018	3,851	69.41	1.40	66.61	72.22		
2018-2019	4,330	67.32	1.39	64.54	70.11		
2019-2020	3,583	71.78	1.41	68.95	74.60		
2020-2021	5,029	66.69	1.37	63.93	69.46		
	σ^2				*SD*		
Random Effects							
Higher Education Study Program	61.37				1,188.10		
Residual	7.83				34.47		

Note. ***p < .001. The dependent variable concerns academic achievement in terms of % study success after the second-chance exam period. The means represent the estimated marginal means (i.e., controlled for higher education study program as a random factor). R_m^2 = marginal R-Squared. R_m^2 represents the proportion of the variance explained solely by the fixed effect. Conditional R-Squared (R_c^2) reflects the proportion of the total variance in the multilevel analysis that is explained by both fixed and random effects. R_c^2 = .05.

Table B4

Pairwise Comparisons Between All Cohorts in Terms of Academic Achievement (i.e., % Study Success After the Second-Chance Exam Period)

Cohort (I)	Cohort (J)	Mean Difference (I-J)	*SE*	95% Confidence Interval		$\lvert \text{Hedges}'\,g \rvert$
				Lower Bound	Upper Bound	
2015-2016	2016-2017	-0.36	0.79	-2.69	1.97	
	2017-2018	-0.46	0.80	-2.81	1.89	
	2018-2019	1.63	0.78	-0.66	3.91	
	2019-2020	-2.82**	0.82	-5.21	-0.43	0.08
	2020-2021	2.26*	0.75	0.05	4.47	0.07
2016-2017	2017-2018	-0.10	0.78	-2.39	2.19	
	2018-2019	1.99	0.76	-0.24	4.22	
	2019-2020	-2.46*	0.80	-4.80	-0.12	0.07
	2020-2021	2.62**	0.73	0.47	4.77	0.08
2017-2018	2018-2019	2.09	0.77	-0.16	4.34	
	2019-2020	-2.36*	0.80	-4.72	-0.01	0.07
	2020-2021	2.72**	0.74	0.55	4.89	0.08
2018-2019	2019-2020	-4.45***	0.78	-6.75	-2.16	0.13
	2020-2021	0.63	0.72	-1.48	2.73	
2019-2020	2020-2021	5.08***	0.76	2.86	7.30	0.15

Note. *$p < .050$, **$p < .010$, ***$p < .001$. The p-values are Bonferroni-adjusted. The mean differences represent the estimated marginal mean differences (i.e., controlled for higher education study program as a random factor).

Table B5

Correlation Matrix (In)Dependent Variable(s)

	1	2	3	4	5	6	7	8	9	10	11	12
1 Study Success After the Second-Chance Exam Period	1											
2 Sex[a]	.09**	1										
3 Socio-Economic Status[b]	-.13**	.03**	1									
4 Education Type Secondary Education[c]	-.21**	-.11**	.10**	1								
5 Hours of Mathematics Secondary Education	.18**	-.23**	-.10*	-.05**	1							
6 Language Proficiency	.12**	-.002	-.09**	-.09**	.04**	1						
7 Self-Control	.12**	.16**	.002	-.01	-.02**	.07**	1					
8 Self-Efficacy (Effort)	.15**	.16**	-.02*	-.004	-.02**	.07**	.51**	1				
9 Self-Efficacy (Comprehension)	.10**	-.14**	-.03**	-.01	.22**	.10**	.24**	.48**	1			
10 Test Anxiety	-.16**	.14**	.04**	.02**	-.10**	-.11**	-.29**	-.23**	-.37**	1		
11 Autonomous Motivation	.10**	.20**	.03**	-.01*	-.04**	.05**	.33**	.47**	.37**	-.11**	1	
12 Controlled Motivation	.03**	-.07**	-.01*	-.02*	.05**	-.04**	-.16**	-.06**	-.02**	.23**	.04**	1

Note. *p < .050, **p < .010. Pearson correlation coefficients are shown for two continuous variables. [a,b,c]Point-biserial correlation coefficients are shown when associated with a continuous variable, and Phi-coefficients when associated with another categorical variable. [a]0 = Male, 1 = Female; [b]0 = High SES, 1 = Low SES; [c]0 = General, 1 = Technical.

Table B6

Overview Included Predictors and Interactions Most Optimal Predictive Model for Academic Achievement (i.e., % Study Success After the Second- Chance Exam Period) (AIC procedure)

	Included in Final Model
Cohort	X
Sex	X
Socio-Economic Status	X
Education Type Secondary Education	X
Hours of Mathematics Secondary Education	X
Language Proficiency	X
Self-Control	X
Self-Efficacy (Effort)	X
Self-Efficacy (Comprehension)	X
Test Anxiety	X
Autonomous Motivation	X
Controlled Motivation	X
Cohort*Sex	-
Cohort*Socio-Economic Status	X
Cohort*Education Type Secondary Education	-
Cohort*Hours of Mathematics Secondary Education	-
Cohort*Language Proficiency	-
Cohort*Self-Control	-
Cohort*Self-Efficacy (Effort)	-
Cohort*Self-Efficacy (Comprehension)	-
Cohort*Test Anxiety	-
Cohort*Autonomous Motivation	-
Cohort*Controlled Motivation	-

Table B7

F-Statistics and Effect Sizes Fixed Effects Multilevel Analysis with Higher Education Study Program as Random Factor

	F	R_m^2	Unique R_m^2
Fixed Effects			
Cohort	10.27***	<.01	<.01
Sex	106.15***	.01	<.01
Socio-Economic Status	188.80***	.01	.01
Education Type Secondary Education	771.01***	.04	.03
Hours of Mathematics Secondary Education	1,138.30***	.08	.06
Language Proficiency	152.97***	.01	.01
Self-Control	19.56***	.02	<.01
Self-Efficacy (Effort)	179.27***	.02	.01
Self-Efficacy (Comprehension)	45.38***	.01	<.01
Test Anxiety	430.25***	.02	.01
Autonomous Motivation	28.84***	.01	<.01
Controlled Motivation	99.94***	<.01	<.01
Cohort*Socio-Economic Status	2.92*	-	-
	σ^2		*SD*
Random Effects			
Higher Education Study Program	86.28		9.29
Residual	1,012.18		31.81

Note. *$p < .050$, **$p < .010$, ***$p < .001$. The dependent variable concerns academic achievement in terms of % study success after the second-chance exam period. Controlled for higher education study program as a random factor. R_m^2 = marginal R-Squared. R_m^2 represents the proportion of the variance explained solely by the fixed effect. Unique R_m^2 indicates the differences between the R_m^2 of the full model and the R_m^2 of the model without a specific fixed factor. Conditional R-Squared (R_c^2) reflects the proportion of the total variance in the multilevel analysis that is explained by both fixed and random effects. R_m^2 full model = .17; R_c^2 full model = .23.

Table B8

Multilevel Analysis Parameter Estimates Fixed Effects with Higher Education Study Program as Random Factor

	B	*SE*	*t*	95% Confidence Interval	
				Lower Bound	Upper Bound
Intercept	-27.67	4.03	-6.86***	-35.55	-19.78
Cohort					
2015-2016	5.24	1.40	3.74***	2.49	7.98
2016-2017	4.99	1.40	3.58***	2.26	7.73
2017-2018	3.78	1.44	2.63**	0.96	6.60
2018-2019	2.91	1.36	2.14*	0.25	5.59
2019-2020	7.29	1.37	5.30***	4.60	9.98
Sex					
Male	-5.12	0.50	-10.30***	-6.10	-4.15
Socio-Economic Status					
High SES	10.23	1.02	10.02***	8.23	12.23
Education Type Secondary Education					
General	23.99	0.86	27.77***	22.29	25.68
Hours of Mathematics Secondary Education	5.45	0.16	33.74***	5.13	5.77
Language Proficiency	1.48	0.12	12.37***	1.24	1.71
Self-Control	0.59	0.13	4.42***	0.33	0.85
Self-Efficacy (Effort)	1.91	0.14	13.38***	1.63	2.18
Self-Efficacy (Comprehension)	-1.08	0.16	-6.74***	-1.40	-0.77
Test Anxiety	-1.99	0.10	-20.74***	-2.18	-1.80
Autonomous Motivation	0.54	0.10	5.37***	0.35	0.74
Controlled Motivation	0.68	0.07	10.00***	0.55	0.82
Cohort*Socio-Economic Status					
2015-2016*High SES	-4.83	1.61	-3.00**	-7.98	-1.67
2016-2017*High SES	-4.07	1.59	-2.55*	-7.19	-0.95
2017-2018*High SES	-2.60	1.63	-1.59	-5.80	0.61
2018-2019*High SES	-4.66	1.56	-2.99**	-7.71	-1.61

	B	*SE*	*t*	95% Confidence Interval	
				Lower Bound	Upper Bound
2019-2020*High SES	-4.28	1.59	-2.68**	-7.40	-1.16

Note. *p < .050, **p < .010, ***p < .001. β = standardized coefficients. The dependent variable concerns academic achievement in terms of % study success after the second-chance exam period. The reference levels for cohort, socio-economic status, sex and education type secondary education are respectively 2020-2021, low SES, female and technical. Controlled for higher education study program as a random factor.

Table B9

*Estimates Cohort*Socio-Economic Status in Terms of Academic Achievement (i.e., % Study Success After the Second-Chance Exam Period)*

Cohort	SES	*n*	*M*	*SE*	95% Confidence Interval	
					Lower Bound	Upper Bound
2015-2016	High SES	2,783	60.64	1.67	57.30	63.98
	Low SES	848	55.24	1.89	51.48	59.00
2016-2017	High SES	3,120	61.15	1.65	57.84	64.46
	Low SES	860	54.99	1.89	51.25	58.74
2017-2018	High SES	3,075	61.41	1.66	58.10	64.73
	Low SES	776	53.78	1.92	49.98	57.58
2018-2019	High SES	3,404	58.49	1.64	55.20	61.79
	Low SES	926	52.92	1.86	49.22	56.62
2019-2020	High SES	2,684	63.24	1.67	59.89	66.59
	Low SES	899	57.29	1.87	53.57	61.01
2020-2021	High SES	3,689	60.23	1.64	56.94	63.51
	Low SES	1,340	50.00	1.77	46.48	53.52

Note. The means represent the estimated marginal means (i.e., controlled for higher education study program as a random factor and for the other appearing predictors in the model). Covariates appearing in the model are evaluated at the following values: Hours of Mathematics Secondary Education = 5.1, Language Proficiency = 17.6, Self-Control = 13.0, Self-Efficacy (Effort) = 15.2, Self-Efficacy (Comprehension) = 14.8, Test Anxiety = 10.0, Autonomous Motivation = 15.0, Controlled Motivation = 8.3.

Table B10

*Pairwise Comparisons Cohort*Socio-Economic Status in Terms of Academic Achievement (i.e., % Study Success After the Second-Chance Exam Period)*

Cohort	SES (I)	SES (J)	Mean Difference (I – J)	*SE*	95% Confidence Interval		$\lvert \text{Hedges}'\, g \rvert$
					Lower Bound	Upper Bound	
2015-2016	High SES	Low SES	5.40**	1.25	1.19	9.62	0.17
2016-2017	High SES	Low SES	6.16***	1.23	2.02	10.23	0.19
2017-2018	High SES	Low SES	7.63***	1.28	3.31	11.95	0.24
2018-2019	High SES	Low SES	5.57***	1.19	1.58	9.56	0.17
2019-2020	High SES	Low SES	5.95***	1.23	1.81	10.09	0.19
2020-2021	High SES	Low SES	10.23***	1.02	6.79	13.67	0.32

Note. $^{**}p < .010$, $^{***}p < .001$. Th *p*-values are Bonferroni-adjusted. The mean differences represent the estimated marginal mean differences (i.e., controlled for higher education study program as a random factor and for the other appearing predictors in the model). Covariates appearing in the model are evaluated at the following values: Hours of Mathematics Secondary Education = 5.1, Language Proficiency = 17.6, Self-Control = 13.0, Self-Efficacy (Effort) = 15.2, Self-Efficacy (Comprehension) = 14.8, Test Anxiety = 10.0, Autonomous Motivation = 15.0, Controlled Motivation = 8.3.

Table B11

Descriptives and Multilevel Analysis Results with Cohort (2015-2016 to 2018-2019, 2019-2020, and 2020-2021) as Fixed Factor and Higher Education Study Program as Random Factor

	n	*M*	*SE*	95% Confidence Interval		*F*	R_m^2
				Lower Bound	Upper Bound		
Fixed Effects							
Cohort						22.64***	.002
2015-2016 to 2018-2019	15,792	68.71	1.31	66.06	71.35		
2019-2020	3,583	71.78	1.41	68.95	74.59		
2020-2021	5,029	66.69	1.37	63.93	69.45		
			σ^2		*SD*		
Random Effects							
Higher Education Study Program			61.15		7.82		
Residual			1,188.45		34.47		

Note. ***p < .001. The dependent variable concerns academic achievement in terms of % study success after the second-chance exam period. The means represent the estimated marginal means (i.e., controlled for higher education study program as a random factor). R_m^2 = marginal R-Squared. R_m^2 represents the proportion of the variance explained solely by the fixed effect. Conditional R-Squared (R_c^2) reflects the proportion of the total variance in the multilevel analysis that is explained by both fixed and random effects. R_c^2 = .05.

Table B12

Pairwise Comparisons Between Cohort 2015-2018, Cohort 2019-2020, and Cohort 2020-2021 in Terms of Academic Achievement (i.e., % Study Success After the Second-Chance Exam Period)

Cohort (I)	Cohort (J)	Mean Difference (I-J)	*SE*	95% Confidence Interval		$\lvert \text{Hedges}' g \rvert$
				Lower Bound	Upper Bound	
2015-2016 to 2018-2019	2019-2020	-3.07***	0.64	-4.60	-1.53	0.09
	2020-2021	2.02**	0.56	0.67	3.36	0.06
2019-2020	2020-2021	5.08***	0.76	3.27	6.89	0.15

Note. $^{**}p < .010$, $^{***}p < .001$. The *p*-values are Bonferroni-adjusted. The mean differences represent the estimated marginal mean differences (i.e., controlled for higher education study program as a random factor).

Table B13

Descriptives and Multilevel Analysis Results Cohort 2020-2021 with SES as Fixed Factor and Higher Education Study Program as Random Factor

	n	*M*	*SE*	95% Confidence Interval		*F*	R^2_m
				Lower Bound	Upper Bound		
Fixed Effects							
SES						165.89***	.03
High SES	3,689	70.84	1.57	67.66	74.03		
Low SES	1,340	56.42	1.75	52.92	59.92		
				σ^2		*SD*	
Random Effects							
Higher Education Study Program				70.06		8.37	
Residual				1,210.03		34.79	

Note. ****p* < .001. The dependent variable concerns academic achievement in terms of % study success after the second-chance exam period. The means represent the estimated marginal means (i.e., controlled for higher education study program as a random factor). The means represent the estimated marginal means (i.e., controlled for higher education study program as a random factor). R^2_m = marginal R-Squared. R^2_m represents the proportion of the variance explained solely by the fixed effect. Conditional R-Squared (R^2_c) reflects the proportion of the total variance in the multilevel analysis that is explained by both fixed and random effects. R^2_c = .08.

Table B14

Descriptives and Multilevel Analysis Results with the (Non-)Cognitive Variables as Dependent Variables, Cohort as Fixed Factor and Higher Education Study Program as Random Factor

	n	*M*	*SE*	95% Confidence Interval		*F*	R_m^2
				Lower Bound	Upper Bound		
Fixed Effects							
Hours of Mathematics Secondary Education						1.38	
2015-2016	3,631	4.97	0.19	4.59	5.34		
2016-2017	3,980	4.96	0.19	4.58	4.34		
2017-2018	3,851	4.97	0.19	4.59	5.35		
2018-2019	4,330	5.02	0.19	4.64	5.40		
2019-2020	3,583	5.01	0.19	4.63	5.39		
2020-2021	5,029	4.99	0.19	4.61	5.37		
Language Proficiency						3.98**	.001
2015-2016	3,631	17.64	0.04	17.55	17.73		
2016-2017	3,980	17.63	0.04	17.55	17.73		
2017-2018	3,851	17.55	0.04	17.54	17.71		
2018-2019	4,330	17.55	0.04	17.47	17.64		
2019-2020	3,583	17.49	0.04	17.40	17.58		
2020-2021	5,029	17.55	0.04	17.46	17.63		
Self-Control						107.75***	.02
2015-2016	3,631	12.60	0.06	12.49	12.72		
2016-2017	3,980	12.65	0.06	12.54	12.76		
2017-2018	3,851	12.67	0.06	12.56	12.78		
2018-2019	4,330	13.39	0.05	13.28	13.50		
2019-2020	3,583	12.94	0.06	12.83	13.05		
2020-2021	5,029	12.95	0.05	12.84	13.05		
Self-Efficacy (Effort)						6.09***	.001
2015-2016	3,631	15.03	0.05	14.93	15.14		
2016-2017	3,980	15.17	0.05	15.07	15.27		

	n	M	SE	95% Confidence Interval		F	R_m^2
				Lower Bound	Upper Bound		
2017-2018	3,851	15.17	0.05	15.07	15.28		
2018-2019	4,330	15.17	0.05	15.07	15.28		
2019-2020	3,583	15.27	0.05	15.17	15.38		
2020-2021	5,029	15.11	0.05	15.01	15.21		
Self-Efficacy (Comprehension)						10.36***	.002
2015-2016	3,631	14.67	0.07	14.52	14.81		
2016-2017	3,980	14.73	0.07	14.59	14.88		
2017-2018	3,851	14.75	0.07	14.61	14.90		
2018-2019	4,330	14.79	0.07	14.65	14.93		
2019-2020	3,583	14.87	0.07	14.73	15.01		
2020-2021	5,029	14.88	0.07	14.74	15.02		
Test Anxiety						0.67	
2015-2016	3,631	9.98	0.07	9.82	10.13		
2016-2017	3,980	10.00	0.07	9.85	10.15		
2017-2018	3,851	10.03	0.07	9.88	10.18		
2018-2019	4,330	9.98	0.07	9.83	10.13		
2019-2020	3,583	10.05	0.07	9.90	10.20		
2020-2021	5,029	10.04	0.07	9.90	10.19		
Autonomous Motivation						3.10**	.001
2015-2016	3,631	15.09	0.09	14.91	15.27		
2016-2017	3,980	15.06	0.09	14.88	15.24		
2017-2018	3,851	15.07	0.09	14.89	15.25		
2018-2019	4,330	14.93	0.09	14.75	15.11		
2019-2020	3,583	15.09	0.09	14.91	15.27		
2020-2021	5,029	15.10	0.09	14.93	15.28		
Controlled Motivation						10.36***	.002
2015-2016	3,631	8.56	0.10	8.37	8.76		
2016-2017	3,980	8.23	0.10	8.04	8.42		
2017-2018	3,851	8.22	0.10	8.03	8.41		

	n	M	SE	95% Confidence Interval		F	R_m^2
				Lower Bound	Upper Bound		
2018-2019	4,330	8.17	0.10	7.98	8.36		
2019-2020	3,583	8.15	0.10	7.95	8.34		
2020-2021	5,029	8.12	0.09	7.93	8.30		

Note. ***p < .001, **p < .010. The means represent the estimated marginal means (i.e., controlled for higher education study program as a random factor). R_m^2 = marginal R-Squared. R_m^2 represents the proportion of the variance explained solely by the fixed effect.

Table B15

Multilevel Analysis Results Random Factor Higher Education Study Program with the (Non-)Cognitive Variables as Dependent Variables and Cohort as Fixed Factor

	σ^2		*SD*	
	Higher Education Study Program	Residual	Higher Education Study Program	Residual
Hours of Mathematics Secondary Education	1.47	1.63	1.21	1.28
Language Proficiency	0.04	2.99	0.20	1.73
Self-Control	0.08	3.40	0.28	1.84
Self-Efficacy (Effort)	0.96	3.73	0.25	1.93
Self-Efficacy (Comprehension)	0.16	2.60	0.41	1.61
Test Anxiety	0.16	5.91	0.40	2.43
Autonomous Motivation	0.24	5.65	0.49	2.38
Controlled Motivation	0.24	9.72	0.49	3.12

Table B16

Pairwise Comparisons Between All Cohorts in Terms of the Included (Non-)Cognitive Variables

	Cohort (I)	Cohort (J)	Mean Difference (I-J)	*SE*	95% Confidence Interval		$\lvert \text{Hedges}' g \rvert$
					Lower Bound	Upper Bound	
Language Proficiency	2015-2016	2016-2017	0.01	0.04	-0.10	0.13	
		2017-2018	0.09	0.04	-0.03	0.21	
		2018-2019	0.09	0.04	-0.02	0.21	
		2019-2020	0.15*	0.04	0.03	0.27	0.09
		2020-2021	0.09	0.04	-0.02	0.21	
	2016-2017	2017-2018	0.07	0.04	-0.04	0.19	
		2018-2019	0.07	0.04	-0.04	0.19	
		2019-2020	0.14*	0.04	0.02	0.25	0.08
		2020-2021	0.08	0.04	-0.03	0.19	
	2017-2018	2018-2019	0.00	0.04	-0.11	0.12	
		2019-2020	0.07	0.04	-0.05	0.18	
		2020-2021	0.00	0.04	-0.10	0.12	
	2018-2019	2019-2020	0.06	0.04	-0.05	0.18	
		2020-2021	0.00	0.04	-0.10	0.11	
	2019-2020	2020-2021	-0.06	0.04	-0.17	0.05	
Self-Control	2015-2016	2016-2017	-0.04	0.04	-0.17	0.08	
		2017-2018	-0.06	0.04	-0.19	0.06	
		2018-2019	-0.79*	0.04	-0.91	-0.67	0.43
		2019-2020	-0.34*	0.04	-0.46	-0.21	0.18
		2020-2021	-0.34*	0.04	-0.46	-0.23	0.19
	2016-2017	2017-2018	-0.02	0.04	-0.14	0.10	
		2018-2019	-0.75*	0.04	-0.87	-0.63	0.40
		2019-2020	-0.29*	0.04	-0.42	-0.17	0.16
		2020-2021	-0.30*	0.04	-0.42	-0.19	0.16
	2017-2018	2018-2019	-0.73*	0.04	-0.85	-0.61	0.39
		2019-2020	-0.27*	0.04	-0.40	-0.15	0.15

	Cohort (I)	Cohort (J)	Mean Difference (I-J)	*SE*	95% Confidence Interval		$\lvert \text{Hedges}'g \rvert$
					Lower Bound	Upper Bound	
		2020-2021	-0.28*	0.04	-0.40	-0.17	0.15
	2018-2019	2019-2020	0.45*	0.04	0.33	0.58	0.25
		2020-2021	0.45*	0.04	0.33	0.56	0.24
	2019-2020	2020-2021	-0.01	0.04	-0.13	0.11	
Self-Efficacy (Effort)	2015-2016	2016-2017	-0.13*	0.04	-0.26	-0.00	0.07
		2017-2018	-0.14*	0.04	-0.27	-0.01	0.07
		2018-2019	-0.14*	0.04	-0.27	-0.01	0.07
		2019-2020	-0.24*	0.05	-0.37	-0.10	0.12
		2020-2021	-0.08	0.04	-0.20	0.05	
	2016-2017	2017-2018	-0.01	0.04	-0.13	0.12	
		2018-2019	-0.01	0.04	-0.13	0.12	
		2019-2020	-0.10	0.04	-0.23	0.03	
		2020-2021	0.06	0.04	-0.06	0.18	
	2017-2018	2018-2019	0.00	0.04	-0.13	0.13	
		2019-2020	-0.10	0.05	-0.23	0.04	
		2020-2021	0.06	0.04	-0.06	0.19	
	2018-2019	2019-2020	-0.10	0.04	-0.23	0.03	
		2020-2021	0.06	0.04	-0.06	0.18	
	2019-2020	2020-2021	0.16*	0.04	0.04	0.28	0.08
Self-Efficacy (Comprehension)	2015-2016	2016-2017	-0.07	0.04	-0.18	0.04	
		2017-2018	-0.09	0.04	-0.20	0.02	
		2018-2019	-0.12*	0.04	-0.23	-0.02	0.08
		2019-2020	-0.20*	0.04	-0.31	-0.09	0.13
		2020-2021	-0.21*	0.04	-0.32	-0.11	0.13
	2016-2017	2017-2018	-0.02	0.04	-0.13	0.09	
		2018-2019	-0.06	0.04	-0.16	0.05	
		2019-2020	-0.14*	0.04	-0.25	-0.03	0.08
		2020-2021	-0.15*	0.03	-0.25	-0.05	0.09
	2017-2018	2018-2019	-0.04	0.04	-0.14	0.07	

	Cohort (I)	Cohort (J)	Mean Difference (I-J)	*SE*	95% Confidence Interval		|Hedges′ g|
					Lower Bound	Upper Bound	
		2019-2020	-0.11*	0.04	-0.22	-0.00	0.07
		2020-2021	-0.13*	0.04	-0.23	-0.02	0.08
	2018-2019	2019-2020	-0.08	0.04	-0.19	0.03	
		2020-2021	-0.09	0.03	-0.19	0.01	
	2019-2020	2020-2021	-0.01	0.03	-0.12	0.09	
Autonomous Motivation	2015-2016	2016-2017	0.03	0.05	-0.13	0.19	
		2017-2018	0.01	0.06	-0.15	0.17	
		2018-2019	0.16	0.05	0.00	0.31	
		2019-2020	-0.00	0.06	-0.17	0.16	
		2020-2021	-0.02	0.05	-0.17	0.14	
	2016-2017	2017-2018	-0.02	0.05	-0.17	0.14	
		2018-2019	0.13	0.05	0.03	0.28	
		2019-2020	-0.03	0.05	-0.19	0.13	
		2020-2021	-0.04	0.05	-0.19	0.10	
	2017-2018	2018-2019	0.14	0.05	-0.01	0.30	
		2019-2020	-0.02	0.05	-0.18	0.15	
		2020-2021	-0.03	0.05	-0.18	0.12	
	2018-2019	2019-2020	-0.16*	0.05	-0.32	-0.00	0.07
		2020-2021	-0.17*	0.05	-0.32	-0.03	0.07
	2019-2020	2020-2021	-0.01	0.05	-0.17	0.14	
Controlled Motivation	2015-2016	2016-2017	0.33*	0.07	0.12	0.54	0.11
		2017-2018	0.34*	0.07	0.13	0.55	0.11
		2018-2019	0.39*	0.07	0.19	0.60	0.13
		2019-2020	0.42*	0.07	0.20	0.63	0.13
		2020-2021	0.45*	0.07	0.25	0.65	0.14
	2016-2017	2017-2018	0.01	0.07	-0.20	0.22	
		2018-2019	0.06	0.07	-0.14	0.26	
		2019-2020	0.09	0.07	-0.13	0.30	
		2020-2021	0.12	0.07	-0.08	0.31	

	Cohort (I)	Cohort (J)	Mean Difference (I-J)	*SE*	95% Confidence Interval		|Hedges′ g|
					Lower Bound	Upper Bound	
	2017-2018	2018-2019	0.05	0.07	-0.15	0.26	
		2019-2020	0.08	0.07	-0.14	0.29	
		2020-2021	0.11	0.07	-0.09	0.30	
	2018-2019	2019-2020	0.02	0.07	-0.18	0.23	
		2020-2021	0.05	0.06	-0.14	0.24	
	2019-2020	2020-2021	0.03	0.07	-0.17	0.23	

Note. $^{*}p < .050$. The p-values are Bonferroni-adjusted. The mean differences represent the estimated marginal means (i.e., controlled for higher education study program as a random factor). Only the pairwise comparisons between the cohorts in terms of the (non-)cognitive variables where cohort has a significant effect are shown (see also Table B14).

3A

Extended Sample Size Data and Results

Table A1

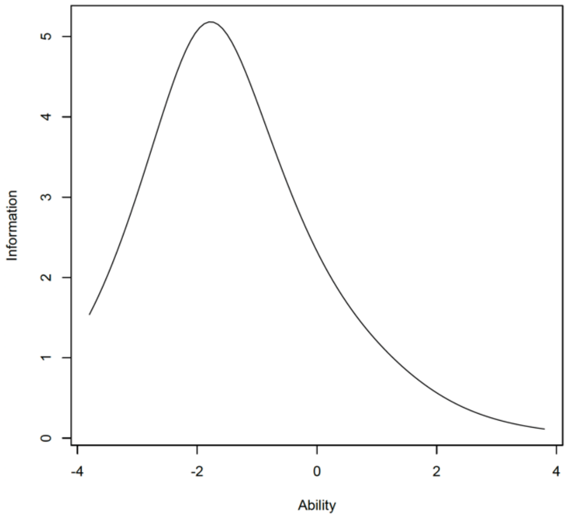
Subsample Sizes per Study Program

	Sex		SES		Total
	Male	Female	Low	High	
Basic mathematics					
Applied Language Studies	105	431	138	398	536
Communication Sciences	110	361	81	390	471
Criminological Sciences	243	928	345	826	1,171
Educational Sciences	33	604	154	483	637
History	242	166	95	313	408
Law	579	1,441	547	1,473	2,020
Pharmaceutical Sciences	220	921	290	851	1,141
Physical Therapy and Motor Rehabilitation	648	1,150	338	1,460	1,798
Political Sciences	258	207	121	344	465
Psychology	395	2,292	769	1,918	2,687
Advanced mathematics					
Biochemistry and Biotechnology	180	250	92	338	430
Biomedical Sciences	235	767	260	742	1,002
Bioscience Engineering	589	601	179	1,011	1,190
Business Administration	938	831	401	1,368	1,769
(Applied/Business) Economics	1,191	761	350	1,602	1,952
Engineering Technology	1,290	168	288	1,170	1,458
Total	7,256	11,879	4,448	14,687	19,135

Note. The advanced mathematics study programs are associated with more advanced mathematical curricula. Through the SIMON project (Fonteyne, 2017), students enrolling in basic mathematics study programs complete a less difficult math assessment compared to students enrolling in more advanced mathematics study programs. For more information on the predictive validity of this basic mathematics test, see Fonteyne and colleagues (2015).

Figure A1

Test Information Curve Three-Parameter Model Sample 2017-2018 to 2019-2020

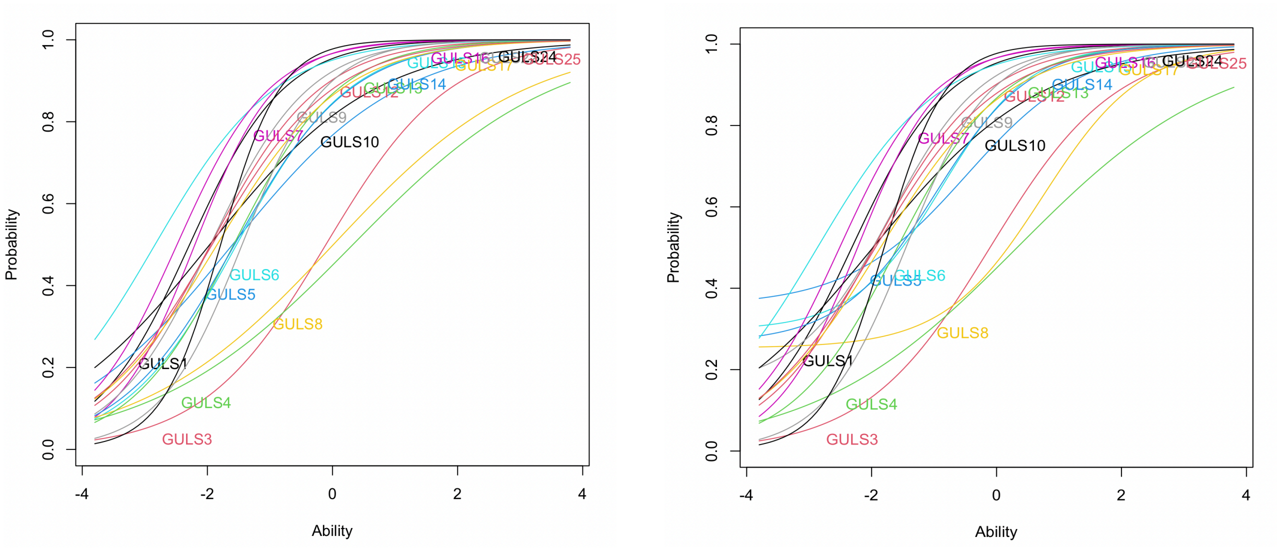


Note. N = 12,527. The three-parameter model offers 82% test information within the lower range of language proficiency ability [-6, 0].

Figure A2

Item Characteristics Curves Two-and Three-Parameter Model Sample 2017-2018 to 2019-2020

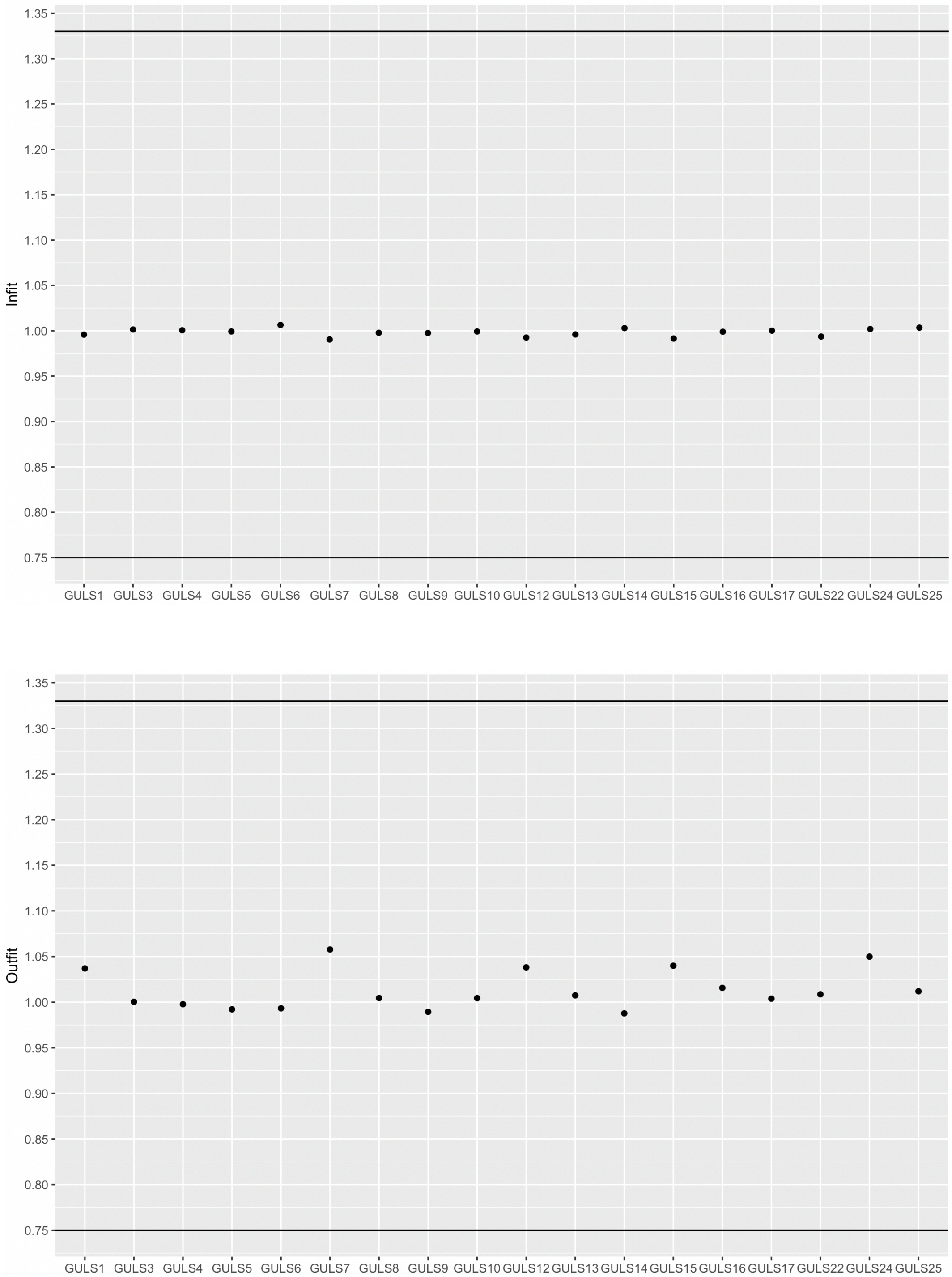
Two-Parameter Model Three-Parameter Model



Note. N = 12,527. Item difficulties range from -2.84 to 0.33, and item discriminations vary between 0.62 and 2.11.

Figure A3

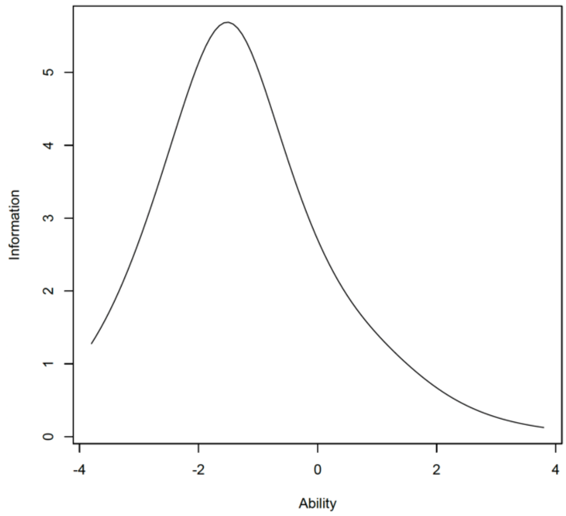
Item Infits (Top) and Outfits (Down) Sample 2017-2018 to 2019-2020



Note. N = 12,527. All item in-and outfit statistics meet the desired criteria of 0.75 to 1.33 (Katz et al., 2021).

Figure A4

Test Information Curve Three-Parameter Model Sample 2020-2021 to 2022-2023

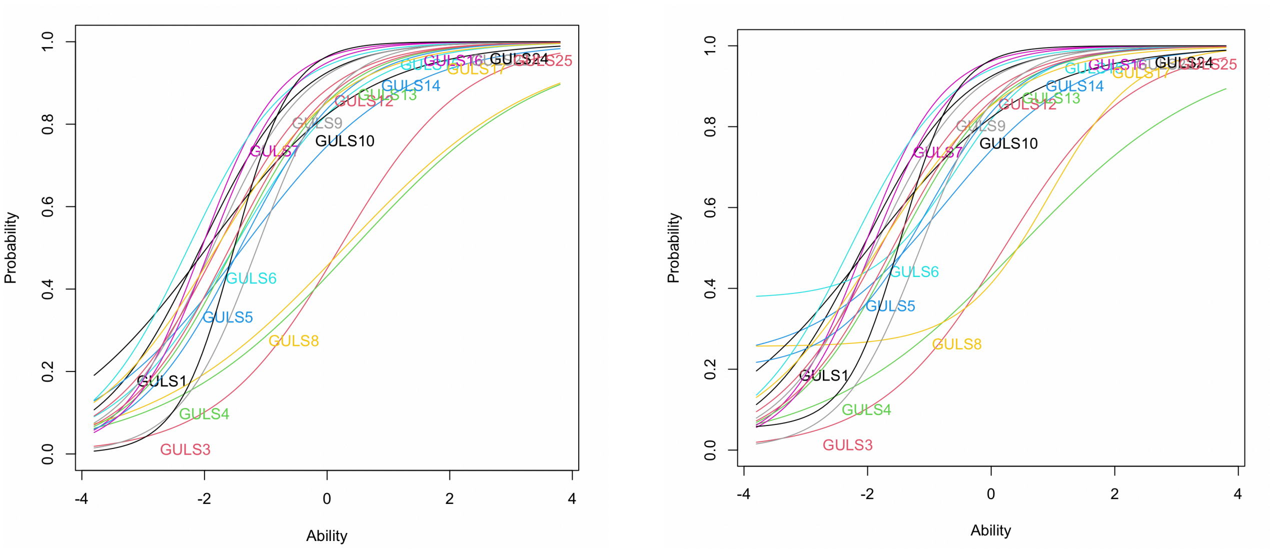


Note. N = 17,204. The three-parameter model offers 81% test information within the lower range of language proficiency ability [-6, 0].

Figure A5

Item Characteristics Curves Two-and Three-Parameter Model Sample 2020-2021 to 2022-2023

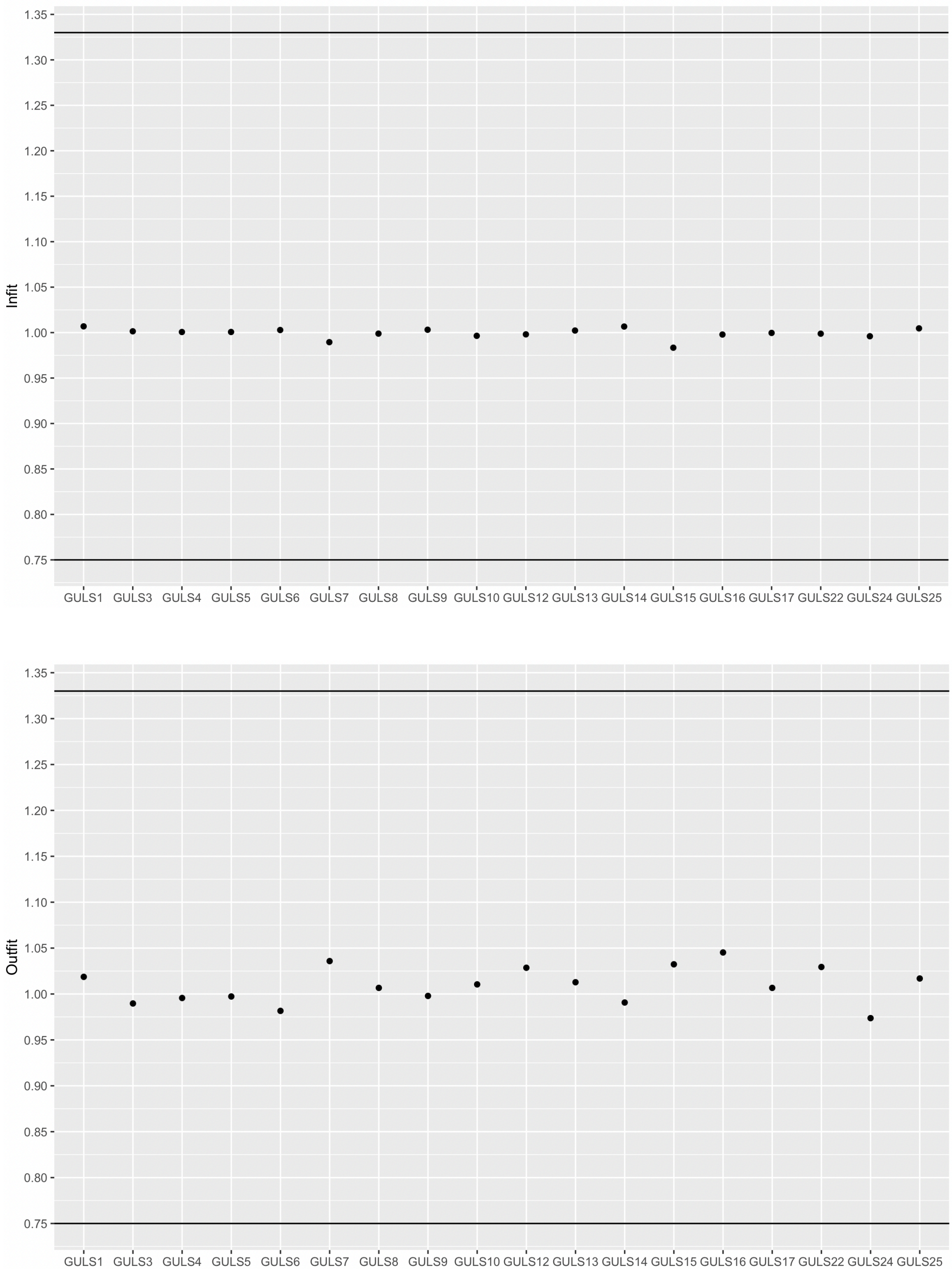
Two-Parameter Model Three-Parameter Model



Note. N = 17,204. Item difficulties span from -2.26 to 0.43, while item discriminations range between 0.62 and 2.17.

Figure A6

Item Infits (Top) and Outfits (Down) Sample 2020-2021 to 2022-2023



Note. $N = 17{,}204$. All item in-and outfit statistics meet the desired criteria of 0.75 to 1.33 (Katz et al., 2021).

Table A2

Correlation Matrix Across and per Subsample

	GPA	Study Success	GULS	Math Hours SE	Vocabulary	SES[a]	Sex[b]	Math Category[c]
GPA	1	.94**	.27**	.25**	.15**	.13**	.08**	.05**
Subsample 1	1	.94**	.24**	.19**	.14**	.09**	.12**	.02
Subsample 2	1	.94**	.28**	.29**	.16**	.17**	.04**	.07**
Study Success	.94**	1	.24**	.23**	.14**	.13**	.06**	.06**
Subsample 1	.94**	1	.22**	.17**	.13**	.09**	.11**	.02
Subsample 2	.94**	1	.26**	.27**	.15**	.17**	.03**	.08**
GULS	.27**	.24**	1	.09**	.32**	.08**	.02*	-.03**
Subsample 1	.24**	.22**	1	.09**	.30**	.06**	-.03**	.00
Subsample 2	.28**	.26**	1	.09**	.33**	.10**	.05**	-.06**
Math Hours SE	.25**	.23**	.09**	1	.05**	.11**	-.18**	.42**
Subsample 1	.19**	.17**	.09**	1	.05**	.11**	-.21**	.43**
Subsample 2	.29**	.27**	.09**	1	.05**	.11**	-.16**	.41**
Vocabulary	.15**	.14**	.32**	.05**	1	.09**	.02*	-.02*
Subsample 1	.14**	.13**	.30**	.05**	1	.09**	-.003	-.002
Subsample 2	.16**	.15**	.33**	.05**	1	.09**	.03**	-.03**

Note. **p* < .05, ***p* < .01. Subsamples 1 (*N* = 8,244) and 2 (*N* = 10,891) respectively encompass students from academic years 2017-2018 to 2019-2020 and 2020-2021 to 2022-2023. [a]Low SES = 0, High SES = 1; [b]Male = 0, Female = 1; [c]Basic mathematics study programs = 0; Advanced mathematics study programs = 1. SE = Secondary Education. The advanced mathematics study programs are associated with more advanced mathematical curricula. Through the SIMON project (Fonteyne, 2017), students enrolling in basic mathematics study programs complete a less difficult math assessment compared to students enrolling in more advanced mathematics study programs. Pearson correlation coefficients are shown for two continuous variables; [a,b,c]Point-biserial correlation coefficients are shown. The advanced mathematics study programs are associated with more advanced mathematical curricula.

References

Fonteyne, L., De Fruyt, F., Dewulf, N., Duyck, W., Erauw, K., Goeminne, K., Lammertyn, J., Marchant, T., Moerkerke, B., Oosterlinck, T., & Rosseel, Y. (2015). Basic mathematics test predicts statistics achievement and overall first year academic success. *European Journal of Psychology of Education*, *30*(1), 95–118. https://doi.org/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. https://doi.org/10.1016/j.lindif.2017.05.003

Katz, D., Clairmont, A., & Wilton, M. (2021). *Measuring what Matters: Introduction to Rasch Analysis in R*. https://bookdown.org/dkatz/Rasch_Biome/

3B

Results Study Success After the Second-Chance Exam Period

Predictive Validity

For the analyses across study programs and study program-specific, no VIF value exceeds 2, indicating the absence of multicollinearity issues.

Across Study Programs

Subsample 1. For the 2017-2018 to 2019-2020 data across 16 study programs, the correlation matrix can be found in Appendix 3A, Table A2. GULS (M = 14.1, SD = 2.9) has a significant effect on study success after the second-chance exam period (hereafter: study success), when controlled for sex, SES, mathematics hours secondary education, vocabulary knowledge, and mathematics category ($t(8{,}237) = 16.55$, $p < .001$, $B = 2.16$, $\beta = 0.18$). Individually, the explained variance of GULS in study success is $R^2 = .05$. GULS's unique contribution to the explained variance in study success is $R^2 = .03$. The difference in explained variance between the full model ($R^2 = .10$) and the model without GULS is significant ($F_{Change}(1, 8{,}237) = 273.80$, $p < .001$). The internal reliability $\omega = .73$.

Subsample 2. For the 2020-2021 to 2022-2023 data across 16 study programs, the correlation matrix is available in Appendix 3A, Table A2. GULS (M = 13.6, SD = 3.1) has a significant effect on study success, when controlled for sex, SES, mathematics hours secondary education, vocabulary knowledge, and mathematics category ($t(10{,}884) = 21.28$, $p < .001$, $B = 2.37$, $\beta = 0.20$). Individually, the explained variance of GULS in study success is $R^2 = .07$. GULS's unique contribution to the explained variance in study success is $R^2 = .04$. The difference in explained variance between the full model ($R^2 = .15$) and the model without GULS is significant ($F_{Change}(1, 10{,}884) = 452.79$, $p < .001$). The internal reliability $\omega = .75$.

Study Program-Specific

The data from 2017-2018 to 2022-2023 shows that GULS significantly predicts study success for all the basic mathematics study programs (M range [13.2, 15.2], SD range [2.2, 3.1]), when controlled for sex, SES, mathematics hours secondary education, vocabulary knowledge and mathematical proficiency (*p*'s < .001). Individually, the explained variance of GULS in study success is $R^2 = .05$ (Physical Therapy and Motor Rehabilitation, and Political Sciences) to $R^2 = .18$ (Applied Language Studies). GULS's unique contribution to the explained variance in study success varies from $R^2 = .01$ (Physical Therapy and Motor Rehabilitation) to $R^2 = .10$ (Applied Language Studies). The differences in explained variance between the full models and the models without GULS are significant (*p*'s < .001).

Across the advanced mathematics study programs (*M* range [12.7, 14.9], *SD* range [2.7, 3.3], the study programs Biochemistry and Biotechnology, and Engineering Technology diverge from the observed pattern where GULS significantly predicts study success, when controlled for sex, SES, mathematics hours secondary education, vocabulary knowledge and mathematical proficiency (*p* range [< .001, .657]). Individually, the explained variance of GULS in study success is $R^2 = .02$ (Biochemistry and Biotechnology, and Engineering Technology) to $R^2 = .06$ (Business Administration, and (Applied/Business) Economics). GULS's unique contribution to the explained variance in study success varies from $R^2 < .01$ (Biochemistry and Biotechnology, and Engineering Technology) to $R^2 = .04$ (Business Administration). The differences in explained variance between the full models and the models without GULS are significant, except for the study programs Biochemistry and Biotechnology, and Engineering Technology (*p* range [< .001, .657]). For details on GULS's descriptives, predictive validity measures for study success, reliability, and the total explained variance of the full model per study program, we refer to Table B1.

Table B1

GULS's Descriptives, Predictive Validity Measures for Study Success, Reliability, and Total Explained Variance Full Model per Study Program

	M (*SD*)	*t*	*B* (*b*)	Individual R^2	Unique R^2	Total R^2	ω
Basic mathematics							
Applied Language Studies	14.7 (2.6)	8.38***	4.83 (0.58)	.18	.10***	.24	.70
Communication Sciences	14.2 (2.6)	4.07***	2.65 (0.65)	.07	.03***	.12	.68
Criminological Sciences	13.5 (2.8)	6.55***	2.65 (0.41)	.07	.03***	.13	.68
Educational Sciences	13.9 (2.6)	5.89***	2.11 (0.36)	.08	.05***	.13	.63
History	15.2 (2.2)	4.08***	3.56 (0.87)	.11	.04***	.15	.64
Law	14.3 (2.7)	12.16***	3.82 (0.31)	.14	.06***	.20	.72
Pharmaceutical Sciences	13.7 (3.1)	5.91***	1.89 (0.32)	.10	.03***	.17	.77
Physical Therapy and Motor Rehabilitation	13.2 (3.1)	5.31***	1.23 (0.23)	.05	.01***	.18	.74
Political Sciences	14.5 (2.5)	3.95***	2.81 (0.71)	.05	.03***	.06	.68
Psychology	13.8 (2.8)	11.33***	2.81 (0.25)	.09	.04***	.18	.69
Advanced mathematics							
Biochemistry and Biotechnology	14.1 (3.1)	1.05	0.64 (0.61)	.02	<.01	.10	.78
Biomedical Sciences	13.7 (3.1)	3.65***	1.45 (0.40)	.05	.01***	.12	.76
Bioscience Engineering	14.9 (2.7)	2.54*	0.97 (0.38)	.05	<.01*	.19	.78
Business Administration	12.7 (3.3)	9.36***	2.20 (0.24)	.06	.04***	.14	.75
(Applied/Business) Economics	13.9 (3.1)	6.75***	1.63 (0.24)	.06	.02***	.13	.78
Engineering Technology	13.4 (3.2)	0.44	0.13 (0.30)	.02	<.01	.13	.77

Note. *p < .05, **p < .01, ***p < .001. ω = McDonald's ω as reliability measure. R^2 = Adjusted R-squared. Total R^2 represents the explained variance by the full model, whereas unique R^2 quantifies the specific contribution of GULS to study success, distinct from other predictors (i.e., sex, SES, mathematics hours secondary education, vocabulary knowledge, and mathematical proficiency). Meanwhile, individual R^2 reflects the overall contribution of GULS to study success, regardless of additional predictors. The advanced mathematics study programs are associated with more advanced mathematical curricula. Through the SIMON project (Fonteyne, 2017), students enrolling in basic mathematics study programs complete a less difficult math assessment compared to students enrolling in more advanced mathematics study programs.

References

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.

3C
Preliminary Analyses Pilot GULS

Participants

Between September 20 and October 6, 2016, $N = 972$ anonymous secondary school students, primarily from the East Flanders region, completed the online GULS with 45 items: $n = 726$ from ASO (general secondary education), $n = 216$ from TSO (technical secondary education), and $n = 30$ from BSO (vocational secondary education).

General Findings

Table C1 presents the mean total scores and standard deviations across the different educational tracks. The variance between these tracks is significant ($F(2,971) = 181.34$, $p < .001$, $R^2 = .27$). Given that the test is not specifically tailored for any track, we will further uniformly analyze GULS.

Table C1

Descriptives Across and per Educational Track

	M	*SD*	*n*
ASO	33.70	5.56	726
TSO	26.05	6.44	216
BSO	22.57	7.23	30
Total	31.66	6.82	972

Note. ASO = general secondary education, TSO = technical secondary education, BSO = vocational secondary education.

Item Analysis

The test is divided into eight categories: Text Structure Recognition (TR - 5 items), Word Meaning (WM - 6 items), Paragraph Construction (PC - 4 items), Text Comprehension (TC - 4 items), Contextual Word Filling (CWF - 6 items), Signal Words (SW - 5 items), Correct Form (CF - 10 items), and Word Combinations (WC - 5 items). Table C2 shows the scoring percentages for each item in descending order.

Table C2

Item Scoring Percentages in Descending Order

Item	*M*	*SD*
WM4	.97	.18
WM2	.96	.19
CF2	.94	.23
WM3	.94	.23
CWF1	.93	.25
CWF5	.92	.27
WM1	.91	.29
WM6	.87	.34
TC4	.87	.34
CF9	.86	.35
WC2	.86	.35
CWF2	.85	.35
CF4	.85	.36
TR2	.84	.37
PC4	.83	.38
CF1	.82	.39
CF3	.81	.39
WC3	.80	.40
TR4	.77	.42
CWF3	.77	.42
SW1	.75	.43
SW5	.75	.43
SW2	.75	.43
WC4	.73	.45
CF8	.72	.45
SW3	.72	.45
PC3	.71	.45
TR1	.70	.46
WC5	.68	.47
TC2	.66	.47
CWF4	.66	.47
SW4	.63	.48
TR5	.63	.48
TR3	.58	.50
PC1	.56	.50
CF5	.53	.50
TC1	.53	.50
PC2	.53	.50
WM5	.43	.50
CF6	.41	.49
CF10	.39	.49
WC1	.37	.48
CWF6	.36	.48
TC3	.34	.47
CF7	.16	.37

Note. $N = 972$. All the items are dichotomous. TR = Text Structure Recognition, WM = Word Meaning, PC = Paragraph Construction, TC = Text Comprehension, CWF = Contextual Word Filling, SW = Signal Words, CF = Correct Form, WC = Word Combinations.

Psychometrics

After analysis, an IRT model on dichotomous data was performed. This model represents the probability of a correct answer given the underlying latent language proficiency of the participant, summarized by the formula:

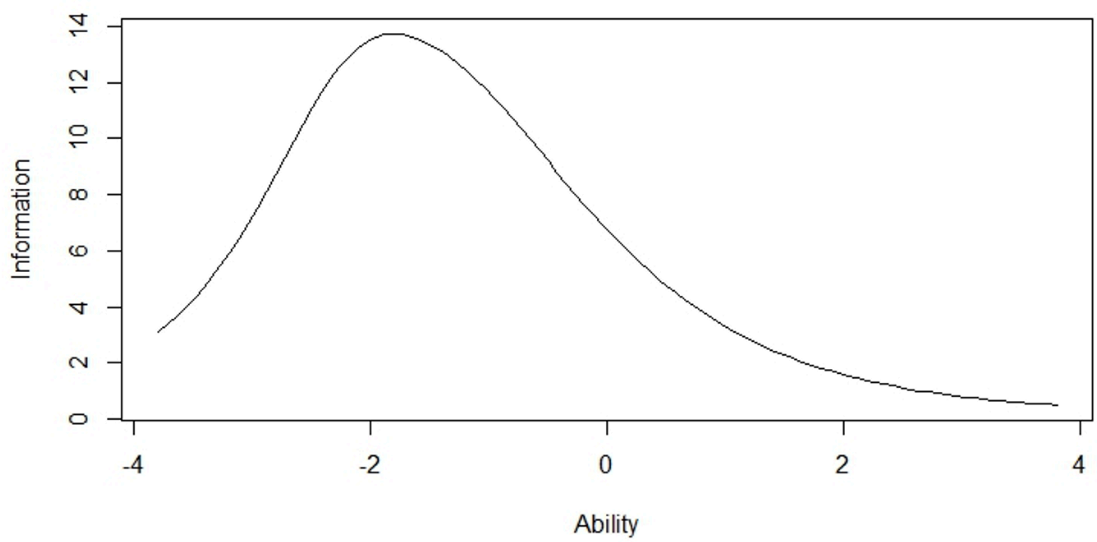
$$P\ (x_{im} = 1 \mid z_m) = g\{\alpha_i(z_m - \beta_i)\}$$

where x_{im} is the score of participant m on item *i*, z_m is the latent proficiency of participant *m*, α_i is the discrimination parameter for item *i* and β_i is the difficulty for item *i*. The logit (or probit) link function is represented by *g*. Implementing a guessing parameter c_i was also explored, which provided a better fit ($p < .001$). Strictly speaking, however, this model could not be applied because some items scored 0 due to blank answers. As the models were similar, only the model without the guessing parameter is discussed further. Nevertheless, it is advisable to consider the model with the guessing parameter in future decisions for a shorter test. The number of blanks is only 0.2% of the total answers.

GULS achieves an internal consistency of Cronbach's $\alpha = 0.85$, considered good to very good in literature (Evers et al., 2009; Nunnally & Bernstein, 1994). Correlations of individual items with the total score (excluding the individual item) range from $r = .11$ (CWF6) to $r = .48$ (PC4), supporting evidence for a common latent construct (language proficiency). Figure 1 shows the test information curve for the model without the guessing parameter. The information in the interval [-6, 0] is 79%, indicating that the test provides information on respondents with lower language proficiency, aligning with the SIMON philosophy. The test is not too difficult, but it is recommended to avoid selecting the easiest items for the final test to maintain sufficient variance. Table C3 shows the item difficulty (*b*) and discrimination (*a*) parameters.

Figure C1

Test Information Curve Two-Parameter Model



Note. The two-parameter model offers 79% test information within the lower range of language proficiency ability [-6, 0].

Table C3

Item Difficulty and Discrimination Parameters

Item	Cluster	*a*	*b*
1	TR1	1.14	-0.92
2	TR2	1.01	-1.90
3	TR3	0.66	-0.51
4	TR4	1.32	-1.23
5	TR5	0.50	-1.13
6	WM1	1.40	-2.10
7	WM2	2.33	-2.22
8	WM3	1.97	-2.12
9	WM4	2.41	-2.22
10	WM5	0.57	0.57
11	WM6	0.90	-2.40
12	PC1	0.57	-0.48
13	PC2	0.65	-0.18
14	PC3	0.71	-1.41
15	PC4	1.85	-1.30
16	TC1	0.88	-0.17
17	TC2	1.32	-0.68
18	TC3	0.70	1.04
19	TC4	1.41	-1.74
20	CWF1	1.50	-2.31
21	CWF2	1.72	-1.50
22	CWF3	1.30	-1.19
23	CWF4	1.00	-0.79
24	CWF5	1.36	-2.33
25	CWF6	0.28	2.12
26	SW1	1.31	-1.13
27	SW2	0.40	-2.82
28	SW3	1.20	-1.01
29	SW4	1.23	-0.58
30	SW5	0.78	-1.62
31	CF1	1.05	-1.72
32	CF2	1.15	-2.92
33	CF3	0.67	-2.37
34	CF4	1.59	-1.51
35	CF5	0.55	-0.25
36	CF6	0.58	0.68
37	CF7	0.42	4.08
38	CF8	1.22	-1.01
39	CF9	1.32	-1.78
40	CF10	0.41	1.10
41	WC1	1.27	0.54
42	WC2	1.40	-1.71
43	WC3	1.53	-1.26
44	WC4	1.25	-1.01
45	WC5	1.41	-0.75

Note. TR = Text Structure Recognition, WM = Word Meaning, PC = Paragraph Construction, TC = Text Comprehension, CWF = Contextual Word Filling, SW = Signal Words, CF = Correct Form, WC = Word Combinations. The item difficulty parameters (*b*) are classified as follows: $b < -2.00$ (very easy), $-2.00 \leq b \leq 2.00$ (moderately difficult), and $b > 2.00$ (very difficult) (Hambleton et al., 1991). The item discrimination parameters (*a*) are categorized in the following way: $0.01 \leq a \leq 0.34$ (very low), $0.35 \leq a \leq 0.64$ (low), $0.65 \leq a \leq 1.34$ (moderate), $1.35 \leq a \leq 1.69$ (high), $a \geq 1.70$ (very high) (Baker, 2001). The items in bold are those that were included in the GULS with 25 items (see also Conclusion and Recommendations).

Conclusion and Recommendations

GULS achieves very good internal consistency and primarily discriminates at the lower end of the latent language proficiency spectrum, in line with the SIMON philosophy. Significant differences between ASO, TSO, and BSO tracks are noted, with lower-scoring tracks exhibiting higher variance, indicating that the language test focuses on the individual rather than the track.

It is recommended not to select the easiest items and to consider the model that corrects for guessing at the item level. The item selection (25 items) to reduce GULS for further validation with first-year higher education students, should consider the item difficulties, discriminatory powers (see bold items in Table C3). We advise to exclude the items related to word meaning in this 25-item GULS, as simple vocabulary knowledge is already assessed in SIMON through the LexTALE test (Lemhöfer & Broersma, 2012).

References

Baker, F. B. (2001). *The Basics of Item Response Theory. Second Edition.* MD: ERIC Clearinghouse on Assessment and Evaluation.

Evers, A., Lucassen, W., Meijer, R., & Sijtsma, A. (2009). *COTAN beoordelingssysteem voor* *de kwaliteit van tests (geheel herziene versie)*. http://www.psynip.nl/website/wat-doethet- nip/tests/beoordelingsprocedure/beoordelingsprocedure

Hambleton, R., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. CA: Sage Publications.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory (3rd ed.)*. New York: McGraw-Hill.

3D
Items GULS

GULS1

Dat de wielersport een omgeving is waarin jargon goed gedijt, is niet verwonderlijk. Dat hebben we in de eerste plaats te danken aan de vroegste wielerverslaggevers, begin twintigste eeuw. Die hadden geen radio en tv, en dus geen geluid of beeld, om de prestaties van de renners te beschrijven. Ze zetten daarom al hun literaire vaardigheden in om de heroïek en dramatiek van de wedstrijden met de pen te beschrijven. In de tweede plaats helpt ook het wielrennen zelf, een sport van grootse tegenstellingen: soms veel tegenslag, soms een beetje geluk, urenlang lijden en kort maar hevig triomferen. Die elementen moeten commentatoren vatten in plastische begrippen zodat wielerliefhebbers de inspanning en de emotie van de renners bijna zelf ervaren. Tot slot is de wielrennerij een vrij besloten wereld: renners, entourage en journalisten leven soms heel dicht op elkaar, dagen- of zelfs wekenlang bij een rittenwedstrijd of grote ronde. In die relatief kleine wereld kan makkelijk een eigen ‘taaltje’ ontstaan.

(Van der Gucht, F., De Caluwe, J., van der Sijs, N. en Janssen, M. (te verschijnen), Atlas van het Nederlands in Vlaanderen [werktitel], Lannoo)

A. De auteur maakt een vergelijking.

B. De auteur onderbouwt een stelling met argumenten.

C. De auteur beschrijft een opeenvolging in de tijd.

D. De auteur stelt een probleem vast en zoekt een oplossing.

GULS2

Bilharzia, of schistosomiasis, is een tropische ziekte die de schistosoma-platwormen veroorzaken. Dat zijn parasieten die in de bloedvaten van de mens leven en zich daar voeden met rode bloedcellen. Ze leggen 300 eitjes per dag die deels via de urine of ontlasting het lichaam verlaten. De eitjes die in het water terechtkomen, ontluiken in larven (miracidia) die een zoetwaterslak als tussengastheer gebruiken. Deze slakken zijn een soort broedkamer waarin de larven zich ontwikkelen en zich ongeslachtelijk voortplanten. De honderdduizenden gekloonde larven (cercaria) verlaten de slak en zetten de jacht in op hun laatste gastheer: de mens. Als de mens zich in besmet water begeeft, dringen de cercaria door de huid het lichaam binnen. Ze scheiden enzymen af die de huid doordringbaar maken. Daarna migreren ze langs het bloedvatenstelsel via de longen en het hart naar de leverpoortader. Van daaruit vertrekken de volwassen wormen naar de bloedvaten rond de urineblaas (urinaire bilharzia) of rond de darmen (intestinale bilharzia), waar ze beginnen aan hun voortplanting. De cyclus begint dan opnieuw.

(http://eoswetenschap.eu/artikel/tropische-ziektes-veroveren-europa)

A. De auteur maakt een vergelijking.

B. De auteur onderbouwt een stelling met argumenten.

C. De auteur beschrijft een opeenvolging in tijd.

D. De auteur stelt een probleem vast en zoekt een oplossing.

GULS3

Zet onderstaande zinnen in de juiste volgorde.

[F] Hoe de dinosauriërs uitstierven, is nog steeds niet helemaal opgehelderd.

[B] Een populaire theorie is dat de meteorietinslag in Mexico een grote hoeveelheid zwavelzuurdeeltjes de lucht in katapulteerde.

[E] Die zouden het zonlicht tegengehouden hebben, zodat het overal ter wereld koud en donker werd.

[C] Die hypothese klopt niet.

[A] Op een ijskoude, donkere aardbol zouden veel meer soorten uitgestorven zijn, zoals de krokodillen.

[D] Bovendien betwijfelen wetenschappers of zwavelzuurdeeltjes wel zo lang in de lucht kunnen blijven na een meteorietinslag.

(http://eoswetenschap.eu/artikel/dinosauri-rs-gingen-rook-op)

GULS4

Zet onderstaande zinnen in de juiste volgorde.

[B] Eicellen worden in het lichaam blootgesteld aan schadelijke afbraakproducten, waaronder vrije radicalen.

[A] Dit zijn reactieve deeltjes die gemakkelijk schade aan kunnen richten in cellen.

[E] Hoe langer een eicel in het lichaam heeft gezeten, hoe meer beschadigingen er zijn opgetreden.

[C] Daardoor raken eicellen die op latere leeftijd vrijkomen minder makkelijk bevrucht.

[D] Ook slagen ze er bijvoorbeeld niet in om zich te nestelen in de baarmoeder waardoor de bevruchting uitloopt op een vroeg miskraam.

(http://www.kennislink.nl/publicaties/hoera-voor-de-oudere-moeder)

GULS5

Resultaten van deze studie toonden aan dat jongere mensen met een hoge socio-economische status (SES) significant meer sporten dan hun oudere metgezellen met een lage SES. Dit is in lijn met ander onderzoek rond kenmerken van sportparticipatie bij de algemene bevolking (Trost et al., 2002). Andere kenmerken zijn het verschil tussen geslacht en etniciteit, waarbij normaal gezien mannen en autochtonen meer sporten dan vrouwen en etnisch-culturele minderheden. Markant in dit onderzoek is dat we geen algemeen verschil zagen tussen mannen en vrouwen, of tussen autochtonen en etnisch-culturele minderheden. Het verschil van geslacht en etniciteit werd volledig verklaard door de lage sportparticipatie van vrouwen van etnisch-culturele afkomst. Deze resultaten zetten nog eens in de verf dat initiatieven, die inzetten op het mobiliseren van deze lage socio-economische groep van etnisch-cultureel diverse afkomst, echt nodig zijn.

(Marlier, M., & Willem, A. (2014). Sport als middel tot integratie van etnisch culturele minderheden en de lagere sociale klasse: resultaten van een studie bij Buurtsport Antwerpen. MOMENTEN (BRUSSEL), (12), 82–86)

Welke stelling kan je uit deze tekst afleiden?

A. Allochtone mannen sporten in het algemeen meer dan vrouwen.
B. In het onderzoek bleek een grote groep allochtone vrouwen weinig te sporten.
C. De studie toont aan dat jonge, laagopgeleide mensen significant meer sporten dan oudere, hoogopgeleiden.

GULS6

Productie van lekkere en gezonde voeding start reeds vanaf de grondstoffen, waarbij wordt gestreefd naar een excellente beginkwaliteit. Tijdens de verwerking wordt meer en meer gebruik gemaakt van mildere conserveringsmethoden (bijvoorbeeld: mildere hittebehandelingen) om deze beginkwaliteit zo goed mogelijk te behouden. In vergelijking met klassieke conserveringsmethoden, resulteert dit echter in vele gevallen in afgewerkte producten met een mindere stabiliteit tijdens bewaring, waarbij zowel microbiologische als chemische afbraakprocessen kunnen optreden. Deze trend naar mildere behandelingen heeft ertoe geleid dat er hogere eisen aan verpakkingen worden gesteld, daar uiteindelijk zij tijdens bewaring deze afbraakprocessen kunnen vertragen.

(*Ragaert, P. (2010). Trends in verpakkingen. FOOD SCIENCE AND LAW, 1(1), 17–20*)

Welke stelling kan je uit deze tekst afleiden?

A. Goede verpakkingen zijn tegenwoordig belangrijker dan de beginkwaliteit van een product.
B. Hoe milder de conserveringsmethoden zijn tijdens de verwerking van voedingsmiddelen, hoe beter de verpakkingen moeten zijn.
C. Milde conserveringsmethoden zoals een mildere hittebehandeling vertragen de afbraakprocessen in voedingsmiddelen.

GULS7

IJzer speelt een belangrijke rol in ons lichaam. Zo zit het in veel eiwitten en is het betrokken bij de opname van zuurstof in ons bloed. Maar als dit metaal in contact komt met UV-straling, wordt het erg instabiel. Dit instabiele ijzer zorgt ervoor dat er zuurstofradicalen ontstaan, die met alles in de cel gaan reageren. Vooral de mitochondriën, de energiecentrales van de cel, hebben hieronder te lijden. Mitochondriën bevatten veel ijzer en zuurstof, dus die gaan al snel kapot door de radicalen. Zonder energiecentrale is de rest van de cel verloren en al snel sterft deze. *(http://www.kennislink.nl/publicaties/zonnebrand-in-de-cel)*

Welke stelling kan je uit deze tekst afleiden?

A. Als het lichaam te veel ijzer opneemt, kunnen lichaamscellen afsterven.
B. Zuurstofradicalen ontstaan als er te veel zuurstof wordt opgenomen in het bloed.
C. UV-straling maakt het ijzer in ons lichaam instabiel.

GULS8

Als ouders hun kinderen steeds minder in dialect opvoeden, waarin voeden ze hun kinderen dan wel op? Uit het onderzoek van Soete (2012) blijkt dat vele ouders hun kinderen in standaardtaal trachten op te voeden, aangezien dat belangrijk zou zijn voor de carrièrekansen van het kind. Dat dat gerapporteerde gedrag ook overeenstemt met het werkelijke gedrag, lijkt echter weinig waarschijnlijk. Een goede standaardtaalbeheersing is voor veel West-Vlamingen immers nog steeds geen evidentie, aldus ook Vandekerckhove (2000). Bovendien staat standaardtaal ook heel ver af van het dialect, de informele omgangstaal waarin de meeste ouders zelf opgevoed zijn.

(*Ghyselen, A.-S. (2012). West-Vlaams is hot, of niet? OVER TAAL (KORTRIJK-HEULE), 51, 101–103*)

Welke stelling kan je uit deze tekst afleiden?

A. Uit het onderzoek van Soete (2012) blijkt dat veel West-Vlaamse ouders hun kinderen in standaardtaal opvoeden.
B. In het onderzoek van Soete (2012) zeiden veel West-Vlaamse ouders dat ze hun kinderen in standaardtaal opvoeden.
C. Dit onderzoek bewijst dat West-Vlaamse ouders vooral dialect spreken met hun kinderen.

GULS9

Welk woord past in beide onderstaande contexten?

A. omkeerbaar
B. irreversibel
C. tijdelijk

- Het aantal zenuwcellen voor een individu ligt bij de geboorte reeds vast voor de rest van het leven. Dit betekent dat beschadigde zenuwcellen zichzelf niet meer kunnen herstellen, en dat die beschadiging dus leidt tot een ……………………………………………… letsel.

- In de fysica spreken we van een ……………………………………………………
proces wanneer we bijvoorbeeld warm water bij koud water voegen. Na een tijdje is de temperatuur van het water overal dezelfde en kunnen we niet meer de ene helft koud en de andere helft warm maken.

GULS10

Welk woord past in beide onderstaande contexten?

A. opponent

B. oppositie

C. tegenligger

- Om in een debat de ander te overtuigen van je gelijk, moet je ervoor zorgen dat je …………………………………………………… je niet in de verdediging drukt of je ter verantwoording oproept.
- Tijdens het EK 2016 hebben de Rode Duivels hun …………………………………………………… Wales erg onderschat. Of dat al dan niet aan de bondscoach lag, laten we liever in het midden.

GULS11

Welk woord past in beide onderstaande contexten?

A. contrasteerde

B. contempleerde

C. contesteerde

- Verwoerd was niet zo tuk op de noemer 'apartheid', hij sprak namelijk liever over 'goed nabuurschap'. Op die manier ……………………………………………… hij de betekenis van wat zich afspeelde in de Zuid-Afrikaanse samenleving.
- De feministe Yosefa Yoteyko ……………………………………………… de gedachte niet dat er meer mannelijke dan vrouwelijke genieën waren – dat gold in de toenmalige wetenschap als een vaststaand feit – maar wel dat daaruit algehele vrouwelijke intellectuele minderwaardigheid zou volgen.

GULS12

Vul het juiste signaalwoord in.

A. **Toch**

B. Bovendien

C. Zo

Een aantal plantensoorten beschikt over aanpassingen die langeafstandsverbreiding mogelijk maken (bijvoorbeeld: windverbreiders). ……………………………………………… zijn de afstanden waarover de meeste soorten verbreiden via zaad beperkt.

(Naar: D'hondt, B., Breyne, P., Van Landuyt, W., & Hoffmann, M. (2012). Draadklaver ontrafeld: de opmerkelijke rol van de mens als verbreider. NATUUR.FOCUS, 11(3), 112–118)

GULS13

Vul het juiste signaalwoord in.

A. daarentegen

B. weliswaar

C. **immers**

Grote mensenmassa's die zich voortbewegen over relatief kleine oppervlaktes vormen voor organisatoren een grote uitdaging. Organisatoren moeten …........................... te allen tijde de veiligheid kunnen garanderen.

(Versichele, M., Neutens, T., Huybrechts, R., Vlassenroot, S., & Gautama, S. (2012). Bluetooth: meer dan gadget voor mobiliteitsonderzoek. VERKEERSSPECIALIST (MECHELEN), (192), 26–29)

GULS14

Duid de juiste omschrijving aan voor het signaalwoord in vet.

A. Op voorwaarde dat

B. Hoewel

C. Zonder dat

Om deze laatste ondervindingen verder te ondersteunen, ging dit onderzoek na of de hypothese ook in de omgekeerde richting opgaat, namelijk een trager herstel van de oogzenuw wanneer er inflammatie gereduceerd wordt. Deze laatste experimenten van de thesis toonden echter geen effect aan. **Mits** er aan de experimentele werkwijze wordt gesleuteld, zou er wel een effect kunnen worden vastgesteld.

(Naar: An Beckers (2005). De invloed van acute inflammatie en inflammaging op het regeneratief potentieel van de zebravisretina, KULeuven)

GULS15

Kies het juiste woord.

Een ……………………………………… is een vermogen dat kennis, inzicht, attitudes en vaardigheden bundelt om in concrete taaksituaties doelen te bereiken.

A. competentie

B. competitie

C. competent

GULS16

Kies het juiste woord.

Ik heb er …………………………………… op aangedrongen bij de directeur om de zaak te herbekijken, maar ze weigert op haar besluit terug te komen.

A. herhaaldelijk

B. herhaald

C. herhalend

GULS17

Kies het juiste woord.

De nieuwe voorzitter moet een …………………………………..………………… reputatie hebben. Als je zo een belangrijke positie bekleedt, kun je geen schandalen gebruiken.

A. omstreden

B. onomstreden

C. onstreden

GULS18

Kies het juiste woord.

Het hof erkent dat dit soort regels nodig kunnen zijn in de praktijk, en dat ze niet altijd in strijd zijn met het verdrag, maar vinden het feit dat geen afwijking mogelijk is ‘buitensporig star en ………………………………………….……… tegenover vrouwen’.

A. discriminatoir

B. discriminatair

C. discriminair

GULS19

Kies het juiste woord.

Het Hof besluit tot een schending van artikel 14 samen met artikel 8, omdat de bepaling van de naam van wettelijke kinderen ‘enkel en alleen gemaakt werd op basis van een discriminatie gebaseerd op het geslacht van de ouders’. Daarnaast erkent het Hof dat dit soort regels gegrond zijn in een ………………………………………………… opvatting van de familie en de macht van de echtgenoot.

A. patriarchische

B. patriarchale

C. patriottische

GULS20

Kies het juiste woord.

Zijn vader was een nogal ……………………………………………………. figuur. Hij had lak aan alle regels en dat bracht hem vaak in conflict met zijn omgeving.

A. inconventioneel

B. onconventioneel

C. aconventioneel

GULS21 tot en met GULS25

Plaats de juiste woordgroep in onderstaande zinnen.

A. staan afwijzend tegenover

B. zijn gebaat bij

C. gaan gepaard met

D. staan haaks op

E. sluiten naadloos aan bij

De ideeën van de communistische partij ……………………………………… (*staan haaks op*) de uitgangspunten van een neoliberaal systeem.

De resultaten van de experimenten uit het onderzoek van professor Vandenberg ……………………………………… (*sluiten naadloos aan bij*) de bevindingen van ons onderzoeksteam. Dat wil zeggen dat beide onderzoeksinitiatieven de hypothese van professor Vandenberg bevestigen.

Drugs ……………………………………… (*gaan gepaard met*) negatieve bijwerkingen.

Zowel automobilisten als fietsers ……………………………………… (*zijn gebaat bij*) goed werkende fietslichten.

Hoewel het wetsvoorstel positief werd onthaald in het parlement, wil de minister het niet aanvaarden. De reden: alle partijleiders ……………………………………… (*staan afwijzend tegenover*) het voorstel.

4A

Extended Sample Size Data

Table A1

Sample Sizes Across and per Academic Year

Academic Year	*n*	Sex		Feedback	
		Male	Female	(Very) Low CoSS	(Fairly) High CoSS
2021-2022	279	89	153	136	106
2022-2023	142	62	88	29	121
Total	392	151	241	165	227

Note. CoSS = Chance of Study Success.

4B
Feedback Examples

Study Program Psychology Very Low Chance of Study Success

SIMON ZEGT

Always bring this personal report to the study or learning path counseling.

Study program: Psychology

Dear First Name,

You recently filled in ‘SIMON zegt’.

SIMON wants to help you and therefore,

- calculates your personal chance of study success in the first year,
- shows you on which skills you score well or not so well and
- indicates what you can do to maximize your skills or where you could still use some support.

Below are your personal results.

Having doubts about your education? Visit your study or learning path counselor.

Would you like to further strengthen certain skills? Be sure to check out the tutoring activities organized by the Monitoring Service and the Study Advice Department.

Do you have questions about ‘SIMON zegt’ feedback? Check out the student guide at ugent.be/simonzegt.

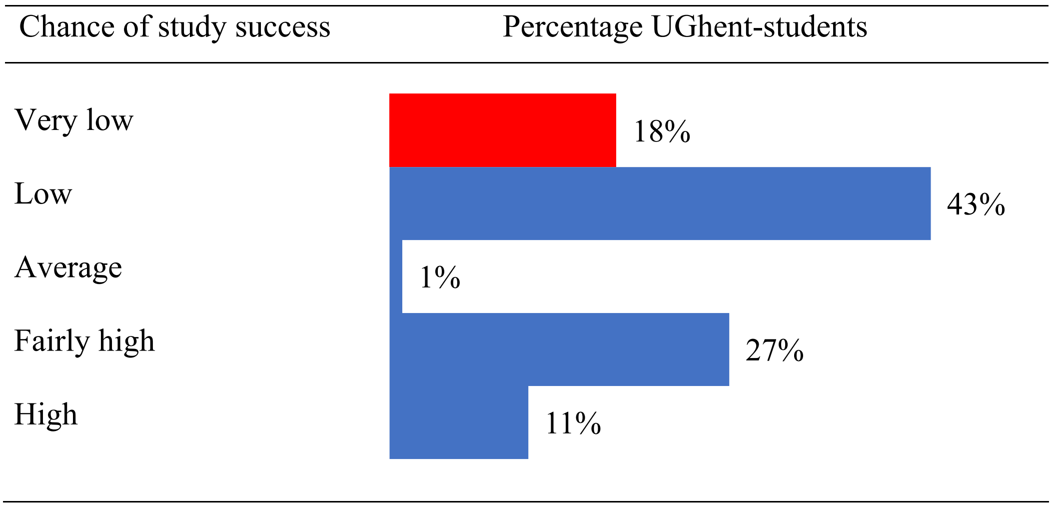
For every issue – no matter how small or ‘innocent’ it may seem – you have somewhere to turn. So, don’t just stick with it alone! Take a look at Feel good about yourself.

Your calculated chance of study success in this study program is VERY LOW

Only 10% of students with this chance of study success achieve the bachelor's degree in this study program within 4 years.

This may come as a bit of a shock. Please know: It is not SIMON's intention to discourage you, but rather just to give you insights into your strengths and possible difficulties when starting higher education.

Below you can see how you (in red) scored, and how other UGhent students from your study program (in blue) scored.



Maybe this study program is not for you at all and you need to rethink your choice of study. Make an appointment and talk about it with a faculty study or learning path counselor or with a student advisor in the Study Advice Department.

Thinking and Reasoning

In the left column, the graph shows how you score **relative to your fellow students in your study program this year.**

In the right column, you will find an **evaluation of your skills** based on research on study success **in your study program**: If you score **low**, you should brush up on this skill in order to tackle your studies in a smooth way. Even if you score **average**, you would do well to strengthen this skill. Below (under 'What now?') you can see how this can be done. If your score is **high**, then in principle you have sufficient skills. You are then of course still welcome at counseling services.

These 2 types of information can sometimes differ from each other. For example, you may score higher than your peers, but still score too 'low' to fully master this skill.

Test, score and position relative to fellow students	Evaluation of your skill
Mathematical reasoning ability (16.80/20)  The middle 50% of the prospective students	Average
Reading comprehension (20/20) This test measures your abilities to process and understand a text.  	High
Vocabulary knowledge (16.25/20)  	Low

Reasoning (14/20) This test measures your reasoning skills. 	Average

What now?

General

View the Montoring Service's offerings of study and learning path guidance here or on the info site of the Monitoring Service Faculty Psychology and Educational Sciences (via Ufora).

Specifically

To tackle your studies in an efficient way, the Monitoring Service organizes the following sessions:

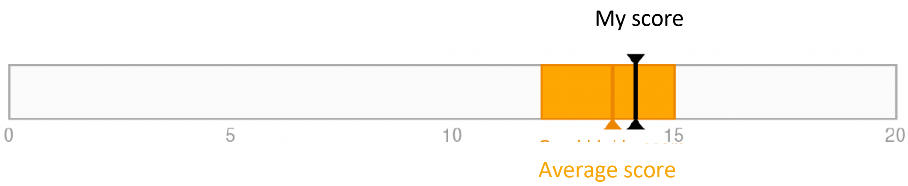
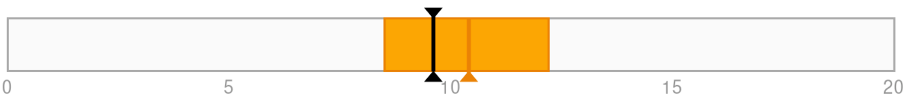
- Study wisely! A 2-part session studying more efficiently for students 1st bachelor Psychology/Pedagogical Sciences (1st and 2nd semester)
- Sessions "How do I study Statistics I?" (1st semester)
- Start2Plan A session study and exam preparation planning for students 1st bachelor Psychology/Pedagogical Sciences (1st and 2nd semester)
- Practice exams (1st semester)

Are you in need of an individual conversation? You can! Make an appointment with one of the study counselors via the e-mail address that belongs to your study program: psychologie.pp@UGent.be or pedawet.pp@UGent.be.

Language coaching and language advice

Looking for **language coaching or language advice**? Take a look at www.ugent.be/taaladvies.

Study Skills

Test, score and position relative to fellow students	Evaluation of your skill
Study skills: knowledge (14.12/20) This test measures your knowledge of study skills and learning strategies.  The middle 50% of the prospective students	Average
Study skills: skills (13.90/20) This test measures your abilities to use learning strategies effectively. 	High
Test anxiety (9.60/20) This test gives a picture of your test anxiety. A high score means more test anxiety. 	High

What now?

General

View the Montoring Service's offerings of study and learning path guidance here or on the info site of the Monitoring Service Faculty Psychology and Educational Sciences (via Ufora).

Specifically

To tackle your studies in an efficient way, the Monitoring Service organizes the following sessions:

- Study wisely! A 2-part session studying more efficiently for students 1st bachelor Psychology/Pedagogical Sciences (1st and 2nd semester)

- Sessions "How do I study Statistics I?" (1st semester)
- Start2Plan A session study and exam preparation planning for students 1st bachelor Psychology/Pedagogical Sciences (1st and 2nd semester)
- Practice exams (1st semester)

Are you in need of an individual conversation? You can! Make an appointment with one of the study counselors via the e-mail address that belongs to your study program: psychologie.pp@UGent.be or pedawet.pp@UGent.be.

High score on test anxiety: take a fear of failure training course.

You will learn to:

- identify and redirect obstructive thoughts
- view situations in a realistic manner
- set achievable goals
- reduce physical tension and stress
- address procrastination and perfectionism

View training offerings.

Personality

<table>
<tr><th>Test, score and position relative to fellow students</th><th>Evaluation of your skill</th></tr>
<tr><td>

Academic self-efficacy: commitment (16/20)

This test gives an idea of how much effort you are willing to put into your studies, even when things are difficult.

The middle 50% of the prospective students

</td><td>Average</td></tr>
<tr><td>

Academic self-efficacy: comprehension (14.86/20)

This test reflects the confidence you have in understanding the subject matter. Too much confidence can have a negative impact on study success.

</td><td>Average</td></tr>
<tr><td>

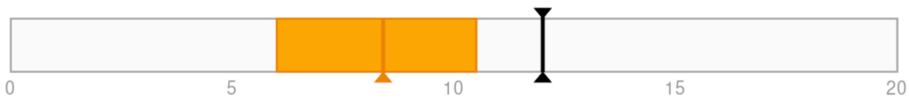
Motivation: autonomous (13/20)

A high score means that you want to take the course primarily because the course interests you and because you think studying is important.

</td><td>Low</td></tr>
<tr><td>

Motivation: controlled (12/20)

A high score means that you want to take the course primarily because others think it is important or because you feel you must take the course.



</td><td>High</td></tr>
<tr><td>

Self-control (16.31/20)

This test measures your self-control of how much you can resist temptation.

</td><td>High</td></tr>
</table>

	
Perseverance (15.33/20) Perseverance is what you need to make a lasting commitment during your studies. 	High

What now?

In a successful study, factors such as motivation, self-confidence, time commitment ... have an important share. This is why we focus on these aspects in the various Monitoring Service's sessions.

Low score on self-control: take a procrastination training course.

You acquire:

- insights into your

procrastination behavior

and learn:

- to handle your study method and schedule efficiently
- self-monitoring techniques
- problem-solving activation
- to clarify postponement excuses

View training offerings.

Summary scores

Below you can find a summary of all the scores achieved.

Section	Subtest	Score (/20)	Evaluation
Thinking and Reasoning	Mathematical reasoning skills	16.80	Average
	Reading comprehension	20	High
	Vocabulary knowledge	16.25	Low
	Reasoning	14	Average
Study Skills	Study skills: knowledge	14.12	Average
	Study skills: skills	13.90	High
	Test anxiety	9.60	High
Personality	Academic self-efficacy: commitment	16	Average
	Academic self-efficacy: understanding	14.86	Average
	Motivation: autonomous	13	Low
	Motivation: controlled	12	High
	Self-control	16.31	High
	Perseverance	15.33	High

Success probability formula

The chance of study success was calculated based on the tests below:

Mathematical reasoning skills; Language Skills; Test Anxiety; Academic self-efficacy: commitment; Academic self-efficacy: effort

For more information on the calculation of the chance of study success, see the guide at www.ugent.be/simonzegt.

Study Program Psychology High Chance of Study Success

SIMON ZEGT

Always bring this personal report to the study or learning path counseling.

Study program: Psychology

Dear First Name,

You recently filled in 'SIMON zegt'.

SIMON wants to help you and therefore,

- calculates your personal chance of study success in the first year,
- shows you on which skills you score well or not so well and
- indicates what you can do to maximize your skills or where you could still use some support.

Below are your personal results.

Having doubts about your education? Visit your study or learning path counselor.

Would you like to further strengthen certain skills? Be sure to check out the tutoring activities organized by the Monitoring Service and the Study Advice Department.

Do you have questions about 'SIMON zegt' feedback? Check out the student guide at ugent.be/simonzegt.

For every issue – no matter how small or 'innocent' it may seem – you have somewhere to turn. So, don't just stick with it alone! Take a look at Feel good about yourself.

Your calculated chance of study success in this study program is HIGH

Students with your scores have a good chance of succeeding in this study program. About 80% of students with a high chance of study success achieve the bachelor's degree in this study program within 4 years. Sustained effort, of course, remains necessary.

Below you can see how you (in green) scored, and how other UGhent students from your study program (in blue) scored.

Chance of study success	Percentage UGhent-students
Very low	18%
Low	43%
Average	1%
Fairly high	27%
High	11%

Would you like to further sharpen certain skills? Be sure to check out the tutoring activities organized by the Monitoring Service and the Study Advice Department.

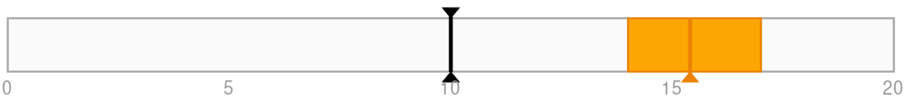
Thinking and Reasoning

In the left column, the graph shows how you score **relative to your fellow students in your study program this year.**

In the right column, you will find an **evaluation of your skills** based on research on study success **in your study program**: If you score **low**, you should brush up on this skill in order to tackle your studies in a smooth way. Even if you score **average**, you would do well to strengthen this skill. Below (under 'What now?') you can see how this can be done. If your score is **high**, then in principle you have sufficient skills. You are then of course still welcome at counseling services.

These 2 types of information can sometimes differ from each other. For example, you may score higher than your peers, but still score too 'low' to fully master this skill.

Test, score and position relative to fellow students	Evaluation of your skill
Mathematical reasoning ability (14.40/20)  The middle 50% of the prospective students	Average
Reading comprehension (12/20) This test measures your abilities to process and understand a text. 	Low
Vocabulary knowledge (16.50/20) 	Low

Reasoning (10/20) This test measures your reasoning skills. 	Low

What now?

General

View the Montoring Service's offerings of study and learning path guidance here or on the info site of the Monitoring Service Faculty Psychology and Educational Sciences (via Ufora).

Specifically

To tackle your studies in an efficient way, the Monitoring Service organizes the following sessions:

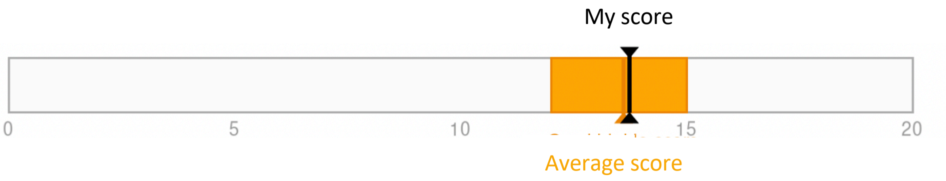
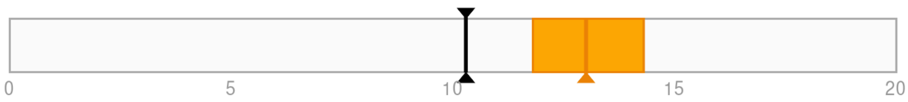
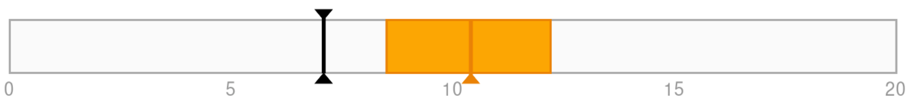
- Study wisely! A 2-part session studying more efficiently for students 1st bachelor Psychology/Pedagogical Sciences (1st and 2nd semester)
- Sessions "How do I study Statistics I?" (1st semester)
- Start2Plan A session study and exam preparation planning for students 1st bachelor Psychology/Pedagogical Sciences (1st and 2nd semester)
- Practice exams (1st semester)

Are you in need of an individual conversation? You can! Make an appointment with one of the study counselors via the e-mail address that belongs to your study program: psychologie.pp@UGent.be or pedawet.pp@UGent.be.

Language coaching and language advice

Looking for **language coaching or language advice**? Take a look at www.ugent.be/taaladvies.

Study Skills

Test, score and position relative to fellow students	Evaluation of your skill
Study skills: knowledge (13.73/20) This test measures your knowledge of study skills and learning strategies.  The middle 50% of the prospective students	Average
Study skills: skills (10.29/20) This test measures your abilities to use learning strategies effectively. 	Low
Test anxiety (7.08/20) This test gives a picture of your test anxiety. A high score means more test anxiety. 	Low

What now?

General

View the Montoring Service's offerings of study and learning path guidance here or on the info site of the Monitoring Service Faculty Psychology and Educational Sciences (via Ufora).

Specifically

To tackle your studies in an efficient way, the Monitoring Service organizes the following sessions:

- Study wisely! A 2-part session studying more efficiently for students 1st bachelor Psychology/Pedagogical Sciences (1st and 2nd semester)
- Sessions "How do I study Statistics I?" (1st semester)

- Start2Plan A session study and exam preparation planning for students 1st bachelor Psychology/Pedagogical Sciences (1st and 2nd semester)
- Practice exams (1st semester)

Are you in need of an individual conversation? You can! Make an appointment with one of the study counselors via the e-mail address that belongs to your study program: psychologie.pp@UGent.be or pedawet.pp@UGent.be.

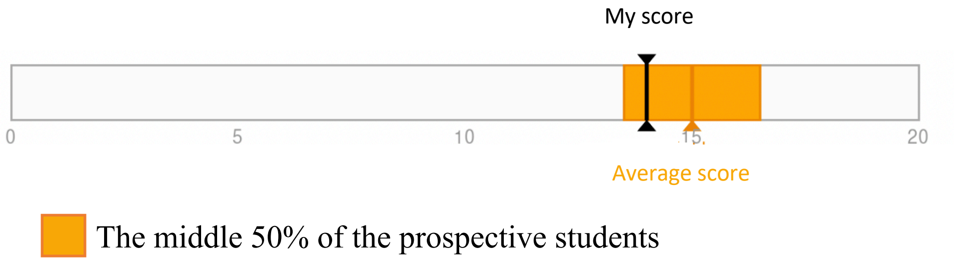
High score on test anxiety: take a fear of failure training course.

You will learn to:

- identify and redirect obstructive thoughts
- view situations in a realistic manner
- set achievable goals
- reduce physical tension and stress
- address procrastination and perfectionism

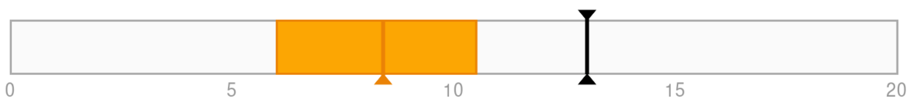
View training offerings.

Personality

<table>
<tr><th>Test, score and position relative to fellow students</th><th>Evaluation of your skill</th></tr>
<tr><td>Academic self-confidence: commitment (14/20)
This test gives an idea of how much effort you are willing to put into your studies, even when things are difficult.


The middle 50% of the prospective students</td><td>Average</td></tr>
<tr><td>Academic self-confidence: comprehension (15.14/20)
This test reflects the confidence you have in understanding the subject matter. Too much confidence can have a negative impact on study success.

</td><td>High</td></tr>
<tr><td>Motivation: autonomous (16/20)
A high score means that you want to take the course primarily because the course interests you and because you think studying is important.

</td><td>Average</td></tr>
<tr><td>Motivation: controlled (13/20)
A high score means that you want to take the course primarily because others think it is important or because you feel you must take the course.

</td><td>High</td></tr>
<tr><td>Self-control (13.54/20)
This test measures your self-control of how much you can resist temptation.</td><td>Average</td></tr>
</table>

<table>
<tr><td></td><td></td></tr>
<tr><td>Perseverance (13.67/20)
Perseverance is what you need to make a lasting commitment during your studies.
</td><td>Average</td></tr>
</table>

What now?

In a successful study, factors such as motivation, self-confidence, time commitment ... have an important share. This is why we focus on these aspects in the various Monitoring Service's sessions.

Low score on self-control: take a procrastination training course.

You acquire:

- insights into your

procrastination behavior

and learn:

- to handle your study method and schedule efficiently
- self-monitoring techniques
- problem-solving activation
- to clarify postponement excuses

View training offerings.

Summary scores

Below you can find a summary of all the scores achieved.

Section	Subtest	Score (/20)	Evaluation
Thinking and Reasoning	Mathematical reasoning skills	14.40	Average
	Reading comprehension	12	Low
	Vocabulary knowledge	16.50	Low
	Reasoning	10	Low
Study Skills	Study skills: knowledge	13.73	Average
	Study skills: skills	10.29	Low
	Test anxiety	7.08	Low
Personality	Academic self-confidence: commitment	14	Average
	Academic self-efficacy: understanding	15.14	High
	Motivation: autonomous	16	Average
	Motivation: controlled	13	High
	Self-control	13.54	Average
	Perseverance	13.67	Average

Success probability formula

The chance of study success was calculated based on the tests below:

Mathematical reasoning skills; Language Skills; Test Anxiety; Academic self-efficacy: commitment; Academic self-efficacy: effort

For more information on the calculation of the chance of study success, see the guide at www.ugent.be/simonzegt.

4C

Items Student Feedback Engagement

Table C1

Items Intentional Feedback Engagement

	Item
IFE1	I will seek more feedback from others to develop competences discussed in my received SIMON feedback
IFE2	I will ask others for suggestions on how I could improve competences described in my received SIMON feedback
IFE3	Through my received SIMON feedback, I will voluntarily participate in study guidance activities organized by Author University
IFE4	Through my received SIMON feedback, I will ask a study counselor, mentor and/or study coach from Author University for a study guidance plan
IFE5	I will search for study guidance activities in line with competences described in my received SIMON feedback
IFE6	Through my received SIMON feedback, I will work on my study behavior

Note. The items were adapted from the behavioral and developmental change dimension of the Feedback in Learning Scale (FLS) (Jellicoe & Forsythe, 2019).

Table C2

Items Behavioral Feedback Engagement

	Item
BFE1	I have sought more feedback from others to develop competences discussed in my received SIMON feedback
BFE2	I have asked others for suggestions on how I could improve competences described in my received SIMON feedback
BFE3	Through my received SIMON feedback, I have voluntarily participated in study guidance activities organized by Author University
BFE4	Through my received SIMON feedback, I have asked a study counselor, mentor and/or study coach from Author University for a study guidance plan
BFE5	I have searched for study guidance activities in line with competences described in my received SIMON feedback
BFE6	Through my received SIMON feedback, I have worked on my study behavior

Note. The items were adapted from the behavioral and developmental change dimension of the Feedback in Learning Scale (FLS) (Jellicoe & Forsythe, 2019).

References

Jellicoe, M., & Forsythe, A. (2019). The Development and Validation of the Feedback in Learning Scale (FLS). *Frontiers in Education*, *4*. https://doi.org/10.3389/feduc.2019.00084

4D
Extended Results

Table D1

Observed Covariance Matrix

	1	2	3	4
1 Feedback with CoSS[a]	0.24			
2 FSE	0.14	0.60		
3 IFE	0.05	0.14	0.76	
4 BFE	0.03	0.02	0.37	0.55

Note. CoSS = Chance of Study Success; FSE = Feedback Self-Efficacy; IFE = Intentional Feedback Engagement; BFE = Behavioral Feedback Engagement. [a]0 = (Very) Low, 1 = (Fairly) High.

Table D2

Model-Implied Covariance Matrix

	1	2	3	4
1 Feedback with CoSS[a]	0.24			
2 FSE	0.14	0.60		
3 IFE	0.05	0.14	0.75	
4 BFE	0.01	0.02	0.37	0.55

Note. CoSS = Chance of Study Success; FSE = Feedback Self-Efficacy; IFE = Intentional Feedback Engagement; BFE = Behavioral Feedback Engagement. [a]0 = (Very) Low, 1 = (Fairly) High.

5A

Extended Sample Size and Study Program Interest Profile Data

Table A1

Sample Sizes of First-Year Students in Academic Years 2016-2017, 2017-2018 and 2018-2019 per Study Program

Study Program	*n*
Applied Language Studies	418
Art History	124
Bio Sciences	189
Biochemistry and Biotechnology	202
Biology	123
Biomedical Sciences	481
Bioscience Engineering	650
Business Administration	868
Communication Sciences	326
Computer Sciences	124
Criminological Sciences	528
Economics, Applied Economics and Business Economics	1,066
Educational Sciences	363
Engineering	728
Engineering - Architecture	179
Engineering Technology	935
History	231
Law	1,017
Linguistics and Literature	495
Medicine	557
Oriental Languages and Cultures	152

Study Program	*n*
Pharmaceutical Sciences	664
Physical Education and Movement Sciences	164
Physical Therapy and Motor Rehabilitation	936
Physics and Astronomy	116
Political Sciences	239
Psychology	1,417
Public Administration and Management	138
Sociology	126
Speech Language and Hearing Sciences	193
Veterinary Medicine	426
Total	14,175

Table A2

Study Program Profiles Correlation Fit and Euclidean Distance Fit, and Sample Sizes of Third-Year Students in Academic Years 2016-2017, 2017-2018 and 2018-2019 per Study Program

Study Program	Profile						*n*
	R	I	A	S	E	C	
Applied Language Studies	8.30	21.67	53.90	44.21	31.92	13.66	164
Art History	17.58	28.71	71.65	36.52	34.49	14.81	47
Bio Sciences	30.00	44.38	18.38	24.96	29.84	23.78	86
Biochemistry and Biotechnology	22.19	53.66	17.98	17.03	15.71	17.25	117
Biology	26.36	54.05	31.45	28.31	21.59	14.82	96
Biomedical Sciences	15.65	54.09	23.25	31.47	21.86	17.15	137
Bioscience Engineering	37.75	51.95	24.50	24.79	33.67	23.80	386
Business Administration	16.95	18.64	24.66	29.45	67.49	56.58	219
Communication Sciences	7.87	22.49	54.28	37.16	54.38	20.99	91
Computer Sciences	32.23	32.12	24.68	9.16	21.63	15.07	29
Criminological Sciences	6.86	29.40	27.34	51.96	35.62	26.32	59
Economics, Applied Economics and Business Economics	23.90	25.67	22.12	24.05	65.06	53.66	442
Educational Sciences	5.97	21.00	31.83	73.40	27.99	13.25	394
Engineering	51.83	40.49	23.87	18.56	41.64	30.02	359
Engineering - Architecture	48.41	37.71	66.47	30.07	35.88	18.03	68
Engineering Technology	52.03	29.43	22.13	16.03	36.31	25.86	268
History	12.55	32.68	52.96	41.54	43.15	21.70	115
Law	11.34	25.97	34.91	45.33	55.09	37.30	80
Linguistics and Literature	8.65	28.42	57.71	46.50	32.37	13.73	197

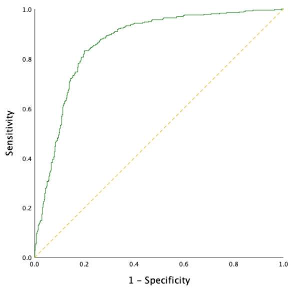
Study Program	Profile						*n*
	R	I	A	S	E	C	
Medicine	17.82	43.71	28.63	50.90	27.13	15.85	322
Oriental Languages and Cultures	14.83	30.61	53.02	46.95	33.00	14.58	34
Pharmaceutical Sciences	13.00	48.42	19.32	36.55	20.90	19.55	353
Physical Education and Movement Sciences	17.44	32.42	21.30	38.96	36.45	21.51	95
Physical Therapy and Motor Rehabilitation	13.09	32.82	21.53	47.45	17.68	11.93	326
Physics and Astronomy	38.38	53.30	23.85	17.95	21.81	19.26	72
Political Sciences	11.86	25.70	35.33	41.92	58.33	33.49	57
Psychology	10.80	31.92	38.84	63.15	34.71	16.86	761
Public Administration and Management	11.71	22.93	30.54	45.20	64.70	39.48	89
Sociology	10.18	34.66	43.80	53.84	37.54	15.58	41
Speech Language and Hearing Sciences	8.30	32.91	34.85	66.56	24.42	15.21	86
Veterinary Medicine	19.50	44.26	25.02	34.85	20.50	17.04	485

5B
Extended Results

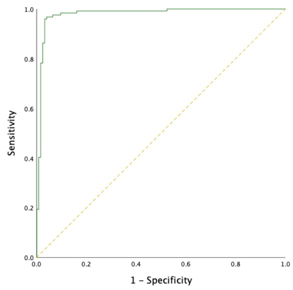
Figure B1

ROC Curves per Study Program

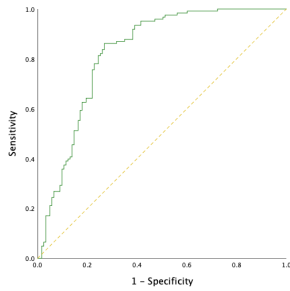
(1) Applied Languages



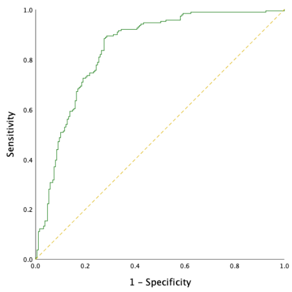
(2) Art History



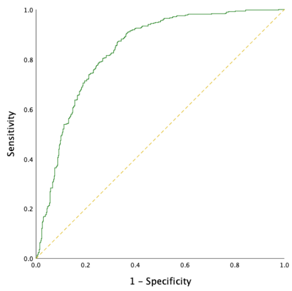
(3) Bio Sciences



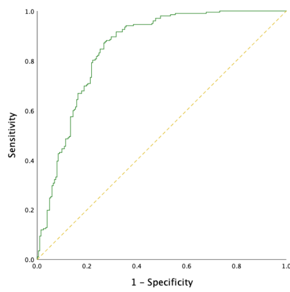
(4) Biochemistry and Biotechnology



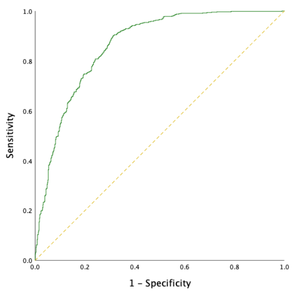
(5) Biology



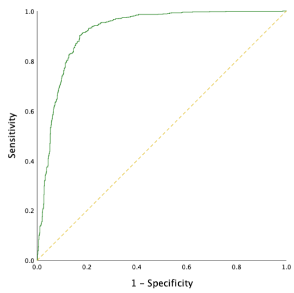
(6) Biomedical Sciences



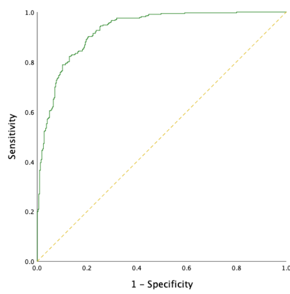
(7) Bioscience Engineering



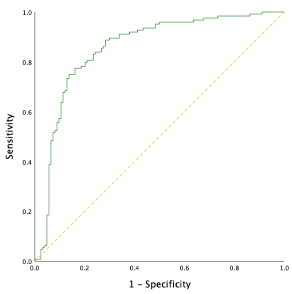
(8) Business Administration



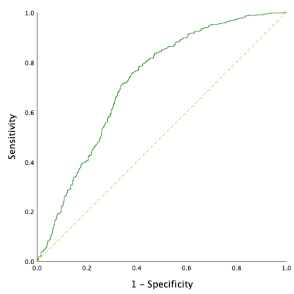
(9) Communication Sciences



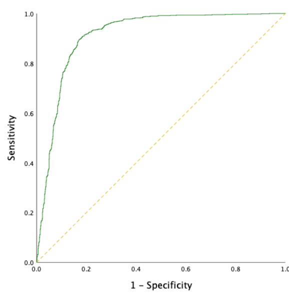
(10) Computer Sciences



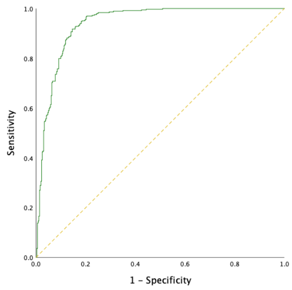
(11) Criminological Sciences



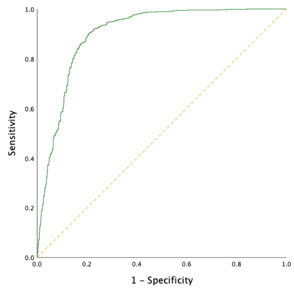
(12) (Applied/Business) Economics



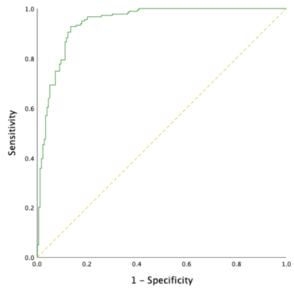
(13) Educational Sciences



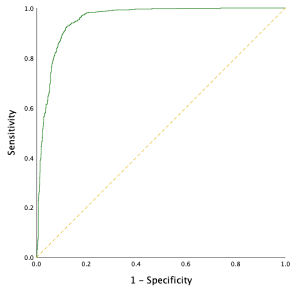
(14) Engineering



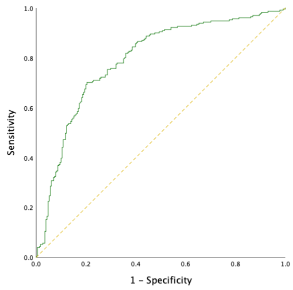
(15) Engineering – Architecture



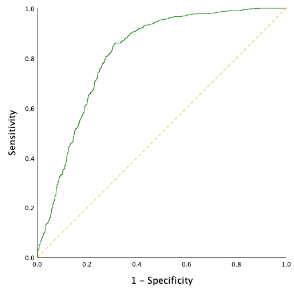
(16) Engineering Technology



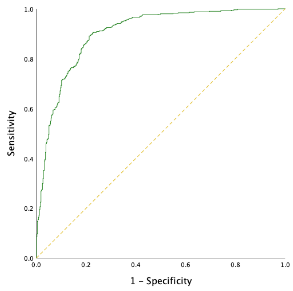
(17) History



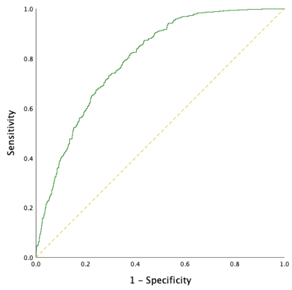
(18) Law



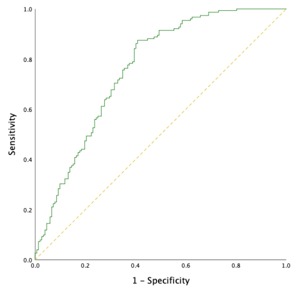
(19) Linguistics and Literature



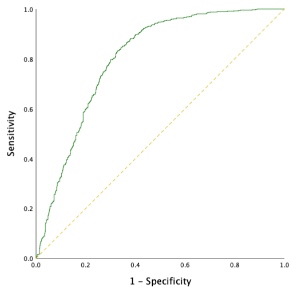
(20) Medicine



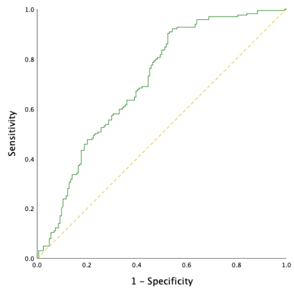
(21) Oriental Languages and Cultures



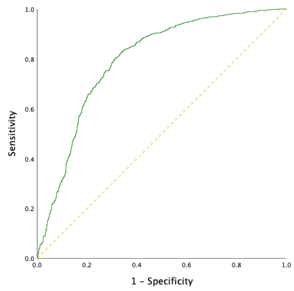
(22) Pharmaceutical Sciences



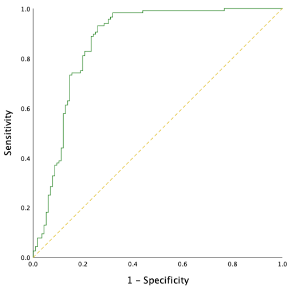
(23) Physical Education and Movement Sciences



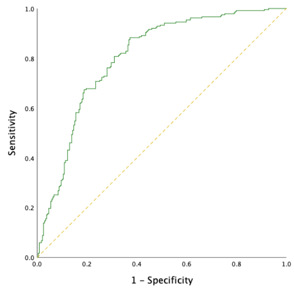
(24) Physical Therapy and Motor Rehabilitation



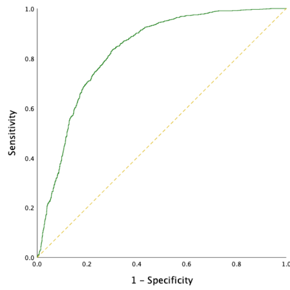
(25) Physics and Astronomy



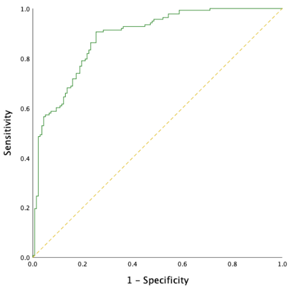
(26) Political Sciences



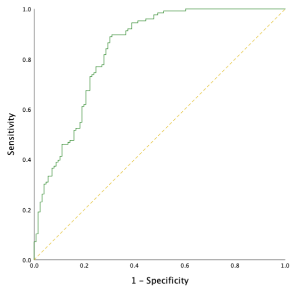
(27) Psychology



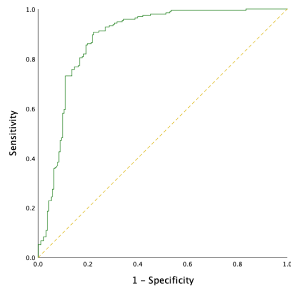
(28) Public Administration and Management



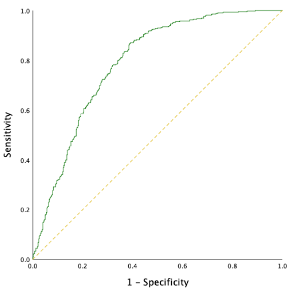
(29) Sociology



(30) Speech language and Hearing Sciences



(31) Veterinary Medicine



Note. Sensitivity refers to the proportion of true positives, or the number of students correctly classified in a specific study program. 1 - Specificity refers to the proportion of false positives, or the number of students incorrectly classified in that specific study program. The green ROC curve delimits The Area Under the Curve (AUC). AUC is a measure of the model's ability to distinguish between students who choose a study program and students who choose another study program, in this case based on PE interest fit (i.e., LRIF). The yellow dotted reference line refers to the 50% probability level benchmark of distinction.

Data Storage Fact Sheets

Chapter 2

Name/identifier study: PhD Mona Bassleer - Chapter 2

Author: Mona Bassleer

Date: 17/06/2024

1. Contact details

1a. Main researcher

- name: Mona Bassleer
- address: Henri Dunantlaan 2 - 9000 Gent - Belgium
- e-mail: mona.bassleer@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Nicolas Dirix (promotor PhD project)
- address: Henri Dunantlaan 2 - 9000 Gent - Belgium
- e-mail: nicolas.dirix@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:

Bassleer, M. (2024). Study orientation in higher education: the Ghent University SIMON project. (Doctoral dissertation). Ghent University, Ghent, Belgium.

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

All data reported in Chapter 2.

3. Information about the files that have been stored

3a. Raw data

----

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

If NO, please justify:

* On which platform are the raw data stored?

 - [X] researcher PC

 - [X] research group file server

 - [X] other (specify): external data drive

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

 - [X] main researcher

 - [X] responsible ZAP

 - [ ] all members of the research group

 - [ ] all members of UGent

- [X] other (specify): members of the SIMON research team

3b. Other files

----

* Which other files have been stored?
  - [ ] file(s) describing the transition from raw data to reported results. Specify: ...
  - [X] file(s) containing processed data. Specify: SPSS files, RStudio 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?
  - [X] individual PC
  - [ ] research group file server
  - [X] other: external data drive

* Who has direct access to these other files (i.e., without intervention of another person)?
  - [X] main researcher
  - [X] 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:

Chapter 3

Name/identifier study: PhD Mona Bassleer - Chapter 3

Author: Mona Bassleer

Date: 17/06/2024

1. Contact details

1a. Main researcher

- name: Mona Bassleer
- address: Henri Dunantlaan 2 - 9000 Gent - Belgium
- e-mail: mona.bassleer@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Nicolas Dirix (promotor PhD project)
- address: Henri Dunantlaan 2 - 9000 Gent - Belgium
- e-mail: nicolas.dirix@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:

Bassleer, M. (2024). Study orientation in higher education: the Ghent University SIMON project. (Doctoral dissertation). Ghent University, Ghent, Belgium.

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

All data reported in Chapter 3.

3. Information about the files that have been stored

3a. Raw data

----

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

If NO, please justify:

* On which platform are the raw data stored?

  - [X] researcher PC

  - [X] research group file server

  - [X] other (specify): external data drive

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

  - [X] main researcher

  - [X] responsible ZAP

  - [ ] all members of the research group

  - [ ] all members of UGent

- [X] other (specify): members of the SIMON research team

3b. Other files

----

* Which other files have been stored?
  - [ ] file(s) describing the transition from raw data to reported results. Specify: ...
  - [X] file(s) containing processed data. Specify: SPSS files, RStudio 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?
  - [X] individual PC
  - [ ] research group file server
  - [X] other: external data drive

* Who has direct access to these other files (i.e., without intervention of another person)?
  - [X] main researcher
  - [X] 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:

Chapter 4

Name/identifier study: PhD Mona Bassleer - Chapter 4

Author: Mona Bassleer

Date: 17/06/2024

1. Contact details

1a. Main researcher

- name: Mona Bassleer

- address: Henri Dunantlaan 2 - 9000 Gent - Belgium

- e-mail: mona.bassleer@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Nicolas Dirix (promotor PhD project)

- address: Henri Dunantlaan 2 - 9000 Gent - Belgium

- e-mail: nicolas.dirix@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:

Bassleer, M. (2024). Study orientation in higher education: the Ghent University SIMON project. (Doctoral dissertation). Ghent University, Ghent, Belgium.

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

All data reported in Chapter 4.

3. Information about the files that have been stored

3a. Raw data

----

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

If NO, please justify:

* On which platform are the raw data stored?

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

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

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

- [X] other (specify): members of the SIMON research team

3b. Other files

----

* Which other files have been stored?
  - [ ] file(s) describing the transition from raw data to reported results. Specify: ...
  - [X] file(s) containing processed data. Specify: SPSS files, RStudio 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?
  - [X] individual PC
  - [ ] research group file server
  - [X] other: external data drive

* Who has direct access to these other files (i.e., without intervention of another person)?
  - [X] main researcher
  - [X] 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:

Chapter 5

Name/identifier study: PhD Mona Bassleer - Chapter 5

Author: Mona Bassleer

Date: 17/06/2024

1. Contact details

1a. Main researcher

\---

- name: Mona Bassleer
- address: Henri Dunantlaan 2 - 9000 Gent - Belgium
- e-mail: mona.bassleer@ugent.be

1b. Responsible Staff Member (ZAP)

\---

- name: Nicolas Dirix (promotor PhD project)
- address: Henri Dunantlaan 2 - 9000 Gent - Belgium
- e-mail: nicolas.dirix@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:

Bassleer, M. (2024). Study orientation in higher education: the Ghent University SIMON project. (Doctoral dissertation). Ghent University, Ghent, Belgium.

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

All data reported in Chapter 5.

3. Information about the files that have been stored

3a. Raw data

----

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

If NO, please justify:

* On which platform are the raw data stored?

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

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

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

- [X] other (specify): members of the SIMON research team

3b. Other files

----

* Which other files have been stored?
  - [ ] file(s) describing the transition from raw data to reported results. Specify: ...
  - [X] file(s) containing processed data. Specify: SPSS files, RStudio 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?
  - [X] individual PC
  - [ ] research group file server
  - [X] other: external data drive

* Who has direct access to these other files (i.e., without intervention of another person)?
  - [X] main researcher
  - [X] 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: