



www.ijemst.net

Students' Perceptions of Low Stakes Positioning Tests at the Start of Higher STEM Education: A Mixed Methods Approach

Jolan Hanssens 
KU Leuven, Belgium

Greet Langie 
KU Leuven, Belgium

Carolien Van Soom 
KU Leuven, Belgium

To cite this article:

Hanssens, J., Langie, G., & Van Soom, C. (2023). Students' perceptions of low stakes positioning tests at the start of higher STEM education: A mixed methods approach. *International Journal of Education in Mathematics, Science, and Technology (IJEMST)*, 11(5), 1094-1112. <https://doi.org/10.46328/ijemst.2889>

The International Journal of Education in Mathematics, Science, and Technology (IJEMST) is a peer-reviewed scholarly online journal. This article may be used for research, teaching, and private study purposes. Authors alone are responsible for the contents of their articles. The journal owns the copyright of the articles. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material. All authors are requested to disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations regarding the submitted work.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Students' Perceptions of Low Stakes Positioning Tests at the Start of Higher STEM Education: A Mixed Methods Approach

Jolan Hanssens, Greet Langie, Carolien Van Soom

Article Info

Article History

Received:

16 September 2022

Accepted:

09 May 2023

Keywords

Students' perception

Perception of assessment

Higher STEM education

Low stakes assessment

Abstract

In Flanders, open admission into higher education has led to heterogeneity in academic preparedness of incoming STEM students. Higher education institutions offer low stakes positioning tests to these students in order to help them assess their level of starting competences. Due to the unique nature of these tests, little can be inferred about students' perception of these tests. In this study, we used mixed methods to investigate these perceptions. Semi-structured interviews (n=20) and thematic analysis revealed that students do not always understand the stakes, objective and structure of positioning tests as intended by the organizers. A questionnaire (n=189) revealed that students considered chemistry problems better reflecting certain study programs than mathematics problems and that they wrongly think chemistry questions are more predictive than mathematics questions. Our findings provide insight in STEM students' multifaceted perceptions of and interaction with low stakes assessment.

Introduction

In Flanders (Belgium) universities have an open admission policy for all incoming students with a Secondary Education (SE) degree into most Higher Education (HE) study programs. Additionally, there is no central standardized test offered at the end of SE. This leads to a heterogeneity in terms of preparedness of incoming students into HE. The under-preparedness of some students can be an issue particularly in those study programs that presuppose a particular level of mathematical competence from the very start, e.g. Science, Technology, Engineering and Mathematics (STEM) programs. Therefore, the Flemish universities organize low stakes positioning tests in order (1) to give students an improved understanding of the expected level of the starting competences and (2) to inform the students about their own level of starting competence compared to that expected level. Students who obtained their SE degree and are considering to enroll into some HE programs can participate in a positioning test during the summer holiday. The content of the positioning tests varies over the study programs (i.e. different study programs have different positioning tests) and over the years.

The nature of positioning tests is low stakes. Their results are non-binding, so low scorers can still enter the study program. In this respect, they clearly differ from more extensively researched types of tests such as central

standardized tests (e.g. Schmitt et al., 2009) and program-specific entrance exams (e.g. Husbands et al., 2015). In such high stakes tests, motivation and test taking behavior could be vastly different (Cole et al., 2008; Eklöf, 2010; Eklöf & Nyroos, 2013; Finn, 2015; Liu et al., 2012). Positioning tests are somewhat similar to placement tests (e.g. Fitchett et al., 2011). The latter guide students to appropriate basic level mathematics courses. They are, however, taken after enrolment in a HE institution and thus lack the opportunity to dissuade students from enrolling or persuade them to take remedial action before they start the study program. So, positioning tests clearly claim a unique position on the international HE landscape.

Since the inception of the first positioning test in 2013, a number of studies have been published on the topic. Most studies have been concerned with the tests' predictive validity. Generally, they found that (1) mathematical tests are better predictors for performance in STEM study programs than other tests like chemistry and academic reading, and (2) that they have incremental predictive validity on top of other predictors like SE academic achievement or learning and study strategies (Fonteyne et al., 2021, 2019; Pinxten et al., 2019; Van den Broeck et al., 2019; Vanderroot et al., 2014, 2015). However, chemistry and academic reading questions have been included in some positioning tests in an attempt to improve participants' perceptions. In terms of consequences, high scores on positioning tests appear to strongly affirm students' study choice, whereas low scores only mildly dissuaded students from starting a certain study program (Fonteyne et al., 2021). In 2019, 68 % of the participants in positioning tests reported to be motivated into remedial action by their participation (Fonteyne et al., 2019). This number increased to 88 % in 2020 (Fonteyne et al., 2021). Quantitative data furthermore shows (1) that most students prepare for the positioning test, (2) that those who prepare generally obtain higher scores and (3) that feedback is perceived generally as useful and fair (Fonteyne et al., 2021, 2019)

Perception and Validity

Validity is a key characteristic of every educational assessment procedure. Does a test measure what it is supposed to measure? seems to be a straightforward question, but various notions of validity exist, some of which are not immediately compatible with one another. Traditional types of test validity are content, construct and criterion based (e.g. predictive) validity. However, Messick (1995) argues that these three are fragmented and that none can really answer the validity question completely. Expanding on traditional validity types, Nevo (1985) operationalizes face validity as perceived or apparent validity of a test. He stresses the importance of face validity in that it may be (1) 'inducing cooperation and positive motivation among subjects before and during the test administration' and (2) 'reducing dissatisfaction and feelings of injustice among low scorers' (Nevo, 1985). Reconciling notions of content, construct, criterion and face validity, Kane (2013) even argues that it is not the test or score in itself that needs validation, but its interpretation and subsequent uses. Consequential validity is therefore a central aspect in building an integrative validity argument (Cook et al., 2015; Kane, 2013; Messick, 1995). In order to understand how students interpret, use and attach consequences to a test, insight in their perception of that test is indispensable. Participants' perceptions can thus be viewed as a necessary part of evidence needed to validate an assessment procedure (Michaelides, 2014). It is furthermore paramount to obtain a complete overview of consequences of test participation, both intended and unintended (Cook et al., 2015; Kane, 2013; Messick, 1995; Michaelides, 2014).

Participants' test perception before, during and after the test, also affects test taking behaviour, which might in turn affect performance and ultimately test validity as well. Perceived importance of a test has been found to increase effort, motivation and performance (Eklöf, 2010; Eklöf & Nyroos, 2013). A widely used framework (e.g. Cole et al., 2008; Knekta & Sundström, 2019) for test motivation is the *expectancy-value theory of achievement motivation* (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000). It states that test taking motivation depends on participants' expectancies for task success and the value they attach to the task of test taking, as composed by one or more of the following value aspects: *importance* (How important is it for a student to engage in the test?), *interest* (Does the student inherently care about or get enjoyment from engaging in the test?), *utility* (Does the test have an apparent practical use, such as entry to a study programme?) and *cost* (How much effort, time and negative emotion does test participation take?) (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000). Specifically, perceptions of the task values *importance* and *utility* have been shown to influence motivation in low stakes tests (Cole et al., 2008). Additionally, it is in such low stakes tests that this link between perception and motivation is especially relevant, as the external motivator associated with higher stakes is missing, which could result into lower motivation and consequently poor test performance (Finn, 2015; Knekta, 2017; Liu et al., 2012). Test anxiety, on the other hand, can have a negative effect on performance (Eum & Rice, 2011). Measuring constructs like motivation or anxiety along with what a test is supposed to measure is of course detrimental for test validity.

Besides validity considerations, it has also been acknowledged that the incorporation of students' perspectives in educational policy and practices is ethically desirable (Elwood, 2012; Elwood & Lundy, 2010; Michaelides, 2014). Students are actors instead of mere subjects in policy enactment and should be treated as such in policy design. They are key stakeholders in educational assessment and therefore, their views should be acknowledged and taken into account. An assessment procedure should ideally be perceived by students as fair, reliable, important, useful, interesting, enjoyable and challenging (Brown & Hirschfeld, 2008; Eccles & Wigfield, 2002; Knekta & Sundström, 2019; Michaelides, 2014; Wigfield & Eccles, 2000).

Aim of This Study

The goal of this study is to evaluate participants' perceptions on positioning tests, as evidence for their validity. As such, this research aims to contribute to filling the knowledge gap around low stakes test perceptions in HE. Additionally, from a practice point of view, increased understanding of students' perceptions can be used to address potential misconceptions by adapting communication and to increase test validity (Cook et al., 2015; Kane, 2013; Messick, 1995; Michaelides, 2014).

Two Positioning Tests

In this study, students' perceptions about the positioning tests for the Bachelor's program in Engineering Technology (ET) and the Bachelor's programs in Chemistry, Biochemistry, Biotechnology, Geography and Geology (CBBGG) are investigated. Both ET and CBBGG have a mathematically focused positioning test, but the study programs clearly comprise more than just mathematics. The inflow of students into these study programs can be considered as heterogeneous, with students typically coming in from low (3 weekly hours of mathematics

class) to advanced (8 hours) level mathematics SE programs (Pinxten et al., 2019). Lastly, efforts to validate the positioning tests for these study programs have been increasingly made in recent years, as their validity is typically lower than that of more mathematically intensive study programs (e.g. Fonteyne et al., 2021; Vanderroost et al., 2015). These efforts include the addition of basic mathematics questions to identify at-risk students (Hanssens et al., 2021a) and experimenting with mathematical context questions to improve perception (Hanssens et al., 2022).

Method

This study presents a mixed method approach to investigating students' perceptions of the positioning tests. On the one hand, thematically analyzed semi-structured interviews were used to gain insight in students' lived experience and on the other hand, a questionnaire was used to compare two test parts in terms of face validity and perceived predictive validity.

Positioning Tests Edition 2020-2021

Participation in the positioning tests of ET and CBBGG was not obligatory in 2020-2021. Furthermore, the positioning tests were transformed to an online format, as opposed to the usual on campus editions of previous years, due to measures in response to the Covid-19 pandemic. This resulted in relatively low participation rates in the official positioning test sessions. In order to obtain more data for validity assessment and to give more incoming freshmen students feedback on their starting level, additional positioning test sessions were organized in the first or second week of the academic year for freshman CBBGG and ET students. Students were informed that participation in these sessions was obligatory for those who had not participated in the official positioning test session. The content of the official and additional session for CBBGG was the same: 10 basic mathematics questions, 15 advanced mathematics questions and 10 chemistry questions. The official test session for ET consisted of the same 10 basic mathematics questions and 15 advanced mathematics questions as the CBBGG test. However, in the additional ET sessions, 10 questions were added compared to the official session. The content of these questions varied: some students got additional mathematics questions without context, others got 10 questions with the same mathematical core but with context added. The purpose of these last 10 questions was to investigate the effect of context on various aspects of the test in a quantitative way, which is outside the scope of this particular study (Hanssens et al., 2022).

Qualitative Interviews

The qualitative aspect of this study is built on 20 semi-structured interviews, with participants being freshmen students both of ET and CBBGG. The interviews were thematically analyzed. In the following paragraphs, the participants, the interview process and the analysis of those interviews are described.

Participants

In the third week of the academic year 2020-2021, an email was sent out to all freshmen CBBGG and ET students

of KU Leuven asking if they wanted to voluntarily participate in an interview about their experience with the positioning test. 20 students responded positively and were interviewed. 11 of them were ET students. Four of those were female and seven were male. Two ET students had participated in the official test session and nine in the additional test session. Nine out of the 20 interviewees were CBBGG students. Five of them were female and four were male. Three participated in the official positioning test session and six in the additional session.

Interview Process

The interviews were conducted during the first semester by the first author via Zoom video calls because of the Covid-19 pandemic. Before each interview, it was specified that the goal of the interview was to gain insight in the student's lived experience and that naturally there were no right or wrong answers. The interview guide is shown in Table 1.

Table 1. Interview Guide

<i>A. Participation</i>
1. How did you feel before the positioning test?
2. What was your general impression after your participation in the positioning test?
3. What do you think of the timing of the test? (Additional session: Imagine you took this test in August, what would you think of that timing?)
<i>B. Opinions about the test</i>
1. What did you think of the difficulty of the questions?
2. What did you think of the content of the questions?
3. Do you have any suggestions to improve the questions?
<i>C. Feedback and consequences</i>
1. How have you experienced the feedback you received?
2. How seriously did you take this feedback?
3. Do you have any suggestions to improve the feedback to the test?
<i>D. Objective of the test</i>
1. What was your goal when participating in the positioning test? (Additional session: What should be the goal of students participating in this test in August?)
2. Did you find the test suitable for your goal?
3. How did the positioning test influence your perception of the study program?
4. Do you think your result on the positioning test can predict your future academic results?

Participants were assured confidentiality in order to avoid socially desirable answers. They were encouraged to reflect critically and honestly so that their responses could be used to improve the positioning test process. After participants gave their consent, the audio of the call was recorded. All interviews lasted between 10 and 25 minutes. After a few practical and introductory questions, the semi-structured interviews consisted of 13 core questions, centered around four themes: (1) *participation*, (2) *opinions about the test*, (3) *feedback and consequences* and (4) *objective of the test*. For each of these core questions, two to five probe questions were

prepared to trigger a response in the case that the participant did not respond adequately to the initial question. Note that questions A3 and D1 were formulated differently for participants in the official and in the additional session (between brackets) of the positioning test.

Analysis

The recorded audio files of these interviews were first transcribed verbatim. These transcripts were then analyzed by means of thematic analysis (Braun & Clarke, 2006). All transcripts were read by the first author once without coding first in order to get familiar with the data. Then, during a second reading, they were coded inductively.

Coding was done based on explicit statements in interview excerpts ranging from one to four sentences. Five transcripts were coded independently by the second and third author and compared to ensure sufficient intercoder reliability. The authors then discussed the codes and organized them into emergent themes and then organized the themes into a thematic map (see Figure 1).

Quantitative Questionnaire

In addition to the interviews, a perception questionnaire was added to the online additional test session for CBBGG students. Two out of the six scales from the job selection questionnaire of Smither et al. (1993) were slightly adapted to fit in our context. The scales used were *face validity* (i.e. how related the test seemed to the study program) and *perceived predictive validity* (i.e. how well the students believed the positioning test could predict future academic performance) with five items each.

All items in these scales were adapted to the specific educational context of this study and translated from English into Dutch by the first author. They were then translated back into English independently by a someone with extensive knowledge of English, in order to verify if the Dutch version did not diverge excessively from the English original. Based on this, a few minor corrections were made in the final version of the questionnaire. An English translation of those final questionnaire items is shown in Table 2.

The items were in the questionnaire twice, immediately after one another, once referring to the mathematics questions and once referring to the chemistry questions. For example, item A1 was in there as ‘I did not understand what the *mathematics* questions had to do with my study program’ and then immediately after as ‘I did not understand what the *chemistry* questions had to do with my study program’. Students were instructed to rate all items on a five point Likert scale (*strongly disagree, disagree, neither disagree nor agree, agree or strongly agree*).

189 out of 221 participants in the additional session of the CBBGG positioning test completed the questionnaire (participation rate: 86 %). Paired two-tailed t-tests were performed in R (v4.0.5) to compare the difference in *face validity* and *perceived predictive validity* between the mathematics and chemistry questions. Bonferroni correction for multiple comparisons was applied.

Table 2. Questionnaire Items

A. Face Validity

1. I did not understand what [the mathematics/ chemistry questions] had to do with my study program.
 2. I could not see any relationship between [the mathematics/ chemistry questions] and what is required in my study program.
 3. It would be obvious to anyone that [the mathematics/ chemistry questions] are related to my study program.
 4. The content of [the mathematics/ chemistry questions] was clearly related to my study program.
 5. There was no real connection between [the mathematics/ chemistry questions] and my study program.
-

B. Perceived predictive validity

1. Failing to pass [the mathematics/ chemistry questions] clearly indicates that you can't do my study program.
 2. I am confident that [the mathematics/ chemistry questions] can predict how well a student will perform in my study program.
 3. My performance on [the mathematics/ chemistry questions] was a good indicator of my ability to do my study program.
 4. It is likely that students who perform well on [the mathematics/ chemistry questions] will perform better in my study program than students who perform poorly.
 5. Teachers can tell a lot about the abilities of students to do my study program from their results on [the mathematics/ chemistry questions].
-

Results

Six distinct themes emerged from the thematic analysis of the 20 semi-structured interviews: (1) *stakes and stress experience*, (2) *objective of participation*, (3) *opinions about the test*, (4) *feedback and consequences of participation*, (5) *perception of predictive validity* and (6) *image of the study programme* (see Figure 1). Note that these themes (Figure 1) are related to the themes in the interview guide (see Table 1), but there are differences, as is typical for thematic analysis (Braun & Clarke, 2006). Most notably, as no participant mentioned predictive validity as an objective of the test, we considered *perception of predictive validity* as a separate theme. Similarly, no participant mentioned influencing their perception of the study program as an objective of the test. This was considered a separate theme as well. Furthermore, due to the nature of responses, the theme *participation* was renamed *stakes and stress experience*. Lastly, the order of the themes was shifted compared to the interview guide in order to obtain a more logically coherent narrative. These six themes and their respective sub-themes will be described in the following paragraphs. Quotes of interviewees will be used as illustration. All of the quotes were translated from Dutch to English by the first author. In order to better frame the quotes, study program (ET or CBBGG), gender and performance on the positioning test (scaled as a mark out of 20, basic math questions not included: 0-7.9/20: low score, 8-13.9/20: medium score, 14-20/20: high score) are reported for each quoted interviewee. The last paragraph of this Results section is dedicated to the qualitative questionnaire data.

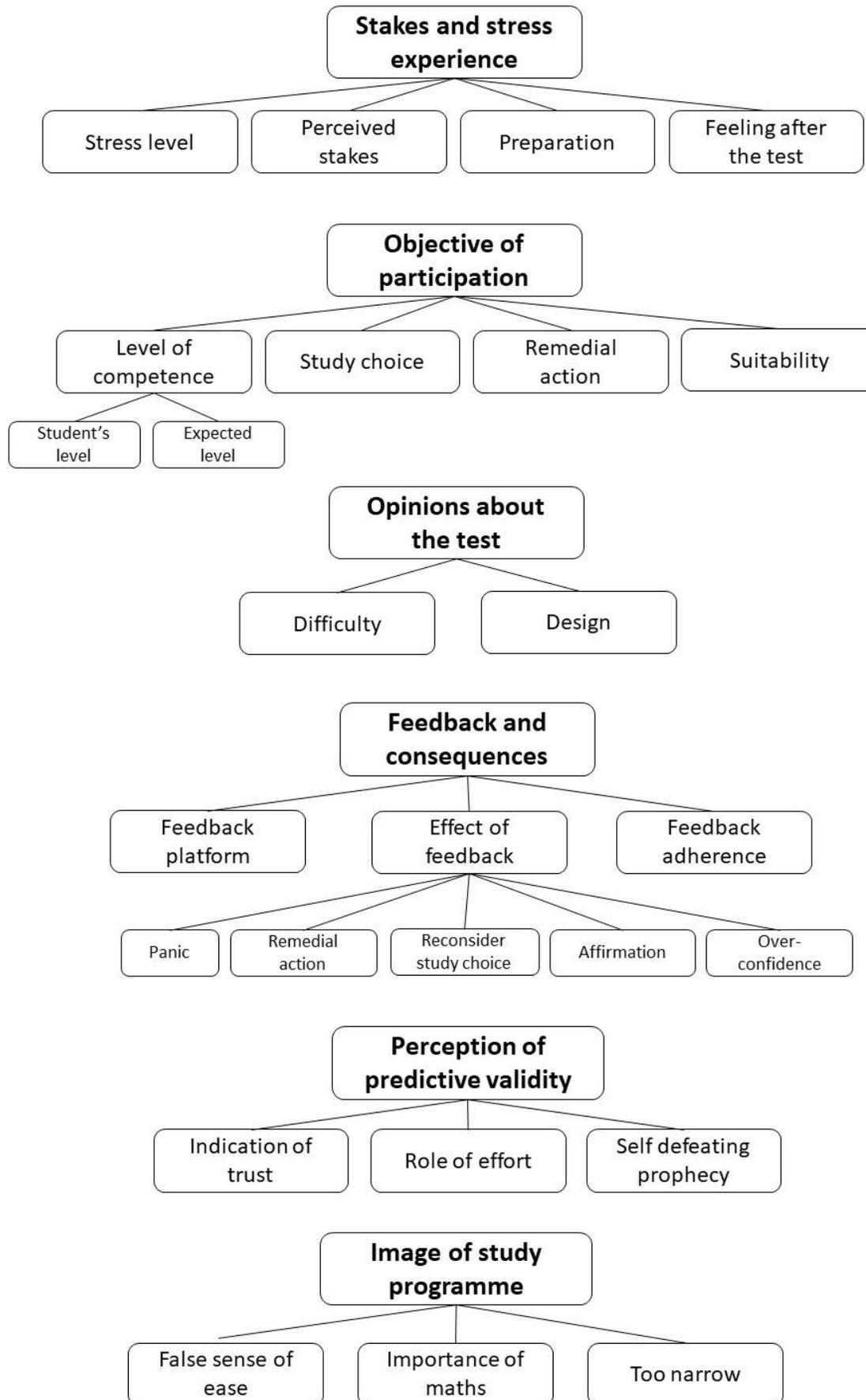


Figure 1. Thematic Map

Stakes and Stress Experience

Stress Level

Most students replied with their stress level after being asked how they felt before the test. Most of them felt no stress: ‘[I felt] chill because I knew it would not count for a grade, so I didn’t have to perform well. Yes, it was just to check my current level, and I thought: “this positioning test is not going to determine whether or not I will pass in the end.”’ (Interviewee 7, CBBGG, female, low score). Others said they felt moderately to highly stressed before the test.

Perceived Stakes

A reason why some students felt stressed, was that they perceived higher stakes than others. While the positioning test is in itself a low stakes test, some students reported that their study choice might in part be determined by it: ‘Fairly stressed out. Because [the positioning test] is a clear sign whether you can do it or not. So, if I had scored badly on it, I would have immediately doubted if this study choice was right for me. So yes, I was really quite stressed.’ (Interviewee 12, ET, male, high score).

Preparation

Around half of the students said they specifically prepared for the positioning test, while the other half reported that they did not. This student gives an interesting justification for not preparing: ‘No, I actually went totally unprepared, because it’s a positioning test and that’s a measure to see what your level is. So, to prepare for that would bias the outcome, I think.’ (Interviewee 6, ET, male, high score).

Feeling after the Test

Most students felt a positive emotion immediately after the test. The most common one was relief. Especially students who reported being stressed before the test, were relieved afterwards: ‘Then I felt normal again. Relieved.’ (Interviewee 12, ET, male, high score). Even those with low scores on the test felt happy it was over. In addition to relief, several students reported a feeling of satisfaction with their performance. Interviewee 18 (CBBGG, male, high score) said he felt anxious for his result and interviewee 8 (CBBGG, male, low score) said he was already disappointed immediately after the test.

Objective of Participation

Level of Competence

Most students said the objective of the positioning test should have a level of competence. They said that the level of competence that they wanted to test is their own, but some students also added that they wanted to test the level of what is expected at the university: ‘Just to check how well a student can do the math and to check what is necessary for the study program. That’s the goal, I think.’ (Interviewee 14, ET, female, medium score).

Study Choice

Some students mentioned that the participation to the positioning test should also assist in making their study choice. Of these students, about half referred to dissuading low scorers and half referred to affirming high scorers. One student mentions both: 'I think motivating your idea of studying [Engineering Technology] if your score's good or else making you really reflect whether it's something for you.' (Interviewee 20, ET, male, high score). Additionally, a few students said that the positioning test could never change their mind about choosing a study program.

Remedial Action

Another commonly mentioned objective for positioning test participation was to motivate lower scoring students to participate in remedial activities: 'If you scored poorly on the positioning test, that might be a trigger to think: "I'll go and take the Summer Course."' (Interviewee 16, ET, male, medium score).

Suitability

A vast majority of students found that the positioning test was suited for their personal objectives, whatever they were. However, one student said that the test was too easy to test her level of competence (Interviewee 4, CBBGG, female, high score).

Opinions about the Test

Difficulty

Most students think that the difficulty level of the questions was balanced: 'I did not think the questions were easy. But, overall, the level is in line with what we're seeing now. I think that the level of the questions is correct.' (Interviewee 5, CBBGG, male, low score). Some others found it rather easy and others again rather difficult.

Design

Most students saw a clear link between the content of the questions on the positioning test and the content of their study program. Especially the mathematical topics on the test and the topics that they were seeing in mathematics class at the time of interviewing seemed to match quite well. Only a few did not see that link. This student gives a remarkable reason for the latter: 'You needed a bit of insight [for the positioning test], but I don't feel like I need to use that same level of insight for the math that we have [in the study program], nor [for] the physics.' (Interviewee 3, ET, male, medium score). Nothing that was on the test, appeared irrelevant, but some students, especially the ones from CBBGG, reported that certain topics were missing. Courses beyond mathematics that were cited to be missing were biology and physics and one student added that statistics questions might be useful: 'A biology part or maybe something like statistics, because that's very important in my study program. And a lot of people find it hard, especially compared to what we did in high school, it's really a lot harder.' (Interviewee 4,

CBBGG, female, high score).

Furthermore, while not all ET students actually had the version of the test with context questions, the ones that remembered that they did, all reported this in a positive way: ‘And those context [questions] are always more fun, I think, because they are linked to something. So you can link it to reality.’ (Interviewee 20, ET, male, high score).

Feedback and Consequences

Feedback Platform

Most students found the feedback platform in which their score and feedback were presented, very clear. Some said they found the comparison with their peers interesting. A few explicitly said they thought the distinction between the basic and advanced mathematics question very interesting, in terms of feedback provided. However, some students would have liked to see feedback per mathematical topic added to the platform and after being asked for suggestions, one of them said model solutions would be useful: ‘Imagine you fail at a type of exercise, exercises about matrices for example. Well, you could put a link to some explanation about how that would work and some examples and so on.’ (Interviewee 20, ET, male, high score).

Effect of Feedback

The interviewees listed five distinct effects that feedback could have on them. First, negative feedback might cause them to panic and stress out. Secondly, on the other hand, most of them said that negative feedback would motivate them into taking remedial action and/or working hard from the very start of their study program. Thirdly, a few said that negative feedback could have caused them to seriously doubt and possibly reconsider their study choice. Then, the most commonly mentioned effect of positive feedback was affirming the students and boosting their confidence. Finally, however, one student said that: ‘Personally, I would not be more motivated. I would think: “if it’s just that...” Then I would really underestimate the math [in the study program], I think.’ (Interviewee 3, ET, male, medium score).

Feedback Adherence

Finally, one student said that low face validity decreases adherence to feedback about reconsidering study choice:

Student: ‘I would be disappointed if my feedback would have been negative. But I would not have reconsidered my study choice. No, I don’t think so. More so because of the fact that there were a lot of very general questions, that were not based on my study program. I would use that as an excuse to feel a bit better.’

Interviewer: ‘Do you think you would take your feedback even more seriously if the test had been more specific to your study program?’

Student: ‘I think so. I mean I might still not reconsider, but in the back of my head, it would more be like: “I failed that test. I will have to do my very best.” ‘ (Interviewee 13, CBBGG, female, medium score).

Perception of Predictive Validity

Indication of Trust

Most students explained their low trust in the predictive validity of the positioning tests by stating that the role of effort and studying was simply much more important than their skills level at the start of the study program: ‘It just depends on how well you study. [...] So you can have a bad positioning test and keep up with maths and study well. Then you’ll have a good exam. It really just depends on how much you do when the time of the exam comes.’ (Interviewee 12, ET, male, high score).

Role of Effort

Most students explained their low trust in the predictive validity of the positioning tests by stating that the role of effort and studying was simply much more important than their skills level at the start of the study program: ‘It just depends on how well you study. [...] So you can have a bad positioning test and keep up with maths and study well. Then you’ll have a good exam. It really just depends on how much you do when the time of the exam comes.’ (Interviewee 12, ET, male, high score).

Self-Defeating Prophecy

Several students mentioned some form of self-defeating prophecy effect, as a type of criterion contamination. This means that their participation and especially their score could affect their academic performance in such a way that low scorers improve their academic performance compared to if they had not participated in the positioning test and high scores could worsen their academic performance. This student explains it very well: ‘However, I think that people who score far above average on their positioning test might start thinking like: “Yeah, I’ve got this. I’ve got my knowledge, so [they] might put in a little less effort.” And that might come back in their face like a boomerang. And it could be vice versa for people who score just below or even far below average. These people will be extra motivated, they will want to prove themselves more and pass as many courses as possible.’ (Interviewee 15, ET, male, medium score). Some students mention just one of these two effects (low scorers performing better or high scorers performing worse than if they had not participated).

Image of the Study Program

A few students said that the positioning test did not affect their image of the study program at all: ‘No, [my image] is still the same. [The positioning test] was never going to change anything. I know what the level is. I also asked around to people from previous year.’ (Interviewee 7, CBBGG, female, low score).

There were also some students to whom the positioning test gave a false sense of ease: ‘Well, if I look at the content we’re seeing in class now, I think those questions were a bit easy. When I saw those questions, I was like: “Well, if this the level of what we’re going to see this year [at university], then things might turn out okay.” But I’m getting anxious again now.’ (Interviewee 18, CBBGG, male, high score).

Another aspect that several students noted is that the test stressed the importance of mathematics in their study program. One ET student mentioned that the context questions were responsible for that: '[...] that you need maths everywhere for problem solving. Practical problems that you'll have to solve by using the theory. That really struck me, especially in that second part [context questions].' (Interviewee 16, male, ET, medium score). On the other hand, some students thought that the mathematical focus of the test gave them a too narrow image of their study program. They were bothered by the fact that other courses from their study program, besides mathematics, were not represented on the positioning test.

Questionnaire – comparing Mathematics and Chemistry Questions

CBBGG students rated the chemistry questions significantly higher than the mathematics questions on both *face validity* and *perceived predictive validity* (Table 3). However, the difference in *perceived predictive validity* is very small.

Table 3. Questionnaire Results

(n=189)	Mathematics (M ± SD)	Chemistry (M ± SD)	P-value (paired t-test, Bonferroni corrected)
<i>Face validity</i>	3.55 ± 0.62	4.04 ± 0.58	2.58e-21
<i>Perceived predictive validity</i>	2.21 ± 0.53	2.29 ± 0.57	0.0135

Discussion and Conclusion

This study sheds light on students' perceptions of two positioning tests in Flanders' open admission educational landscape. It is the first study that uses qualitative and quantitative methods to describe perceptions on Flemish positioning tests in detail. It contributes to the understanding of students' perceptions of low stakes tests, and has direct implications for practice and policy on the local level.

Students' Perceptions of the Positioning Test

Some students report higher experienced stakes than others. While it was the aim of this study to evaluate the perception of a low stakes test, the universal low stakes character of the test is thus challenged. It is wrong to assume that, because admission to HE does not formally depend on the result, positioning tests are consequently experienced as low stakes assessment by every single student. After all, the study choice of some students may depend on it, at least in part. Whether a student is susceptible to being influenced in their study choice, could depend on which stage of their decision making process they are in. Students can start to commit to a specific study choice before they are enrolled, but this is not always the case (Germeijs & Verschueren, 2006). The higher stakes experienced by some students can be related to the importance as well as the utility task value of expectancy-value theory (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000). The utility in this case lies in guiding students to the right choice whether to enrol or not. Importantly, only some students reported a high stakes

experience, while most still considered the positioning test low stakes. This confirms previous research that found that the actual proportion of students that could get dissuaded from their study choice by a positioning test, is rather small (Fonteyne et al., 2021, 2019). The difference in experienced stakes between participants complicates interpretation of results and validation of the tests (Cook et al., 2015; Kane, 2013). Higher experienced stakes could lead to increased motivation but higher anxiety as well (Eklöf, 2010; Eum & Rice, 2011; Finn, 2015; Knekta, 2017; Reich et al., 2015; Simzar et al., 2015). This is illustrated by the abovementioned quote from interviewee 12 (ET, male, high score) as well: ‘Fairly stressed out. Because [the positioning test] is a clear sign whether you can do it or not. So, if I had scored badly on it, I would have immediately doubted if this study choice was right for me. So yes, I was really quite stressed.’ While the remarkable stakes character of Flemish positioning tests highlights their uniqueness, it is possible that in other situations of assumed low stakes educational assessment, a similar difference in experienced stakes exists. It is important that teachers, test designers and researchers be aware of this possibility.

Most students said they wanted to test their level of competence with the positioning test, which is also confirmed quantitatively by Fonteyne et al. (2021). The subsequent use of this information, seemed to depend on their score: students said the lowest scorers should maybe reconsider, fairly low scorers should take remedial action and the highest scorers should get some affirmation. This is very similar to what is mentioned on the website of the positioning tests (ijkingsstoets.be). Of course, students could have their image of the positioning test shaped by official communication like this. Note that no student mentioned all three of these actions, though. All the mentioned objectives can be considered as utility value aspects of expectancy-value theory (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000), which concurs with earlier research that found that utility is one of the main value aspects in a low stakes test (Cole et al., 2008).

Remarkably, most ET students thought the content of the test questions adequately reflected their expectation of the study program, while many CBBGG students noted that they missed certain topics like biology, physics and statistics. ET students thus generally perceived a higher content validity than CBBGG students. A possible explanation for this difference between is that the ET test consisted only of mathematics, which may have shaped the students’ perception that the positioning test is a mathematics test. On the other hand, the CBBGG test had chemistry questions as well, possibly providing the image of a broader test. Within the mathematical part of the test, ET and CBBGG students generally reported a high perception of content validity (i.e. no relevant topics were missing).

Both in the questionnaire as in the interviews, students reported moderate to high face validity. CBBGG students reported chemistry questions seeming more related to their study program than mathematics questions. This could be attributed to the non-program specific nature of mathematics. From the interviews, context questions on the ET positioning test also seemed to increase face validity. One student said that increased face validity would lead to increased feedback adherence, which is in line with Nevo (1985).

Participants’ low perception of predictive validity can weaken the validity argument of the positioning test (Cook

et al., 2015; Kane, 2013). Predictive validity has been the main research topic related to positioning tests so far (Vanderoost et al., 2014, 2015, Pinxten et al., 2019, Van den Broeck et al., 2019, Fonteyne et al., 2019, 2021, Hanssens et al., 2021a, 2021b) and it is also a key characteristic in the design of the tests. The assumption of students' trust in the predictive validity is also key in the validity argument. Information about the predictive value of the test and thus importance of mathematics as a necessary starting competence for both study programs is incorporated in the feedback to convince students to critically reflect on their level of preparedness and, in the case of a low score, potentially reconsider their study choice. A lack of trust in predictive validity thus weakens the implications the test can have, and therefore weakens the entire validity argument.

The most common argument students give for their low perception of predictive validity, is their trust in their ability to affect their academic performance by studying adequately. This type of growth mindset has been shown to drastically increase learning outcomes (Yeager et al., 2019). However, students do not seem to understand that the predictive value of the positioning test and of studying adequately are not necessarily mutually exclusive (Pinxten et al., 2019). Interestingly, CBBGG students even rated chemistry questions a little more predictive than mathematics questions on the questionnaire, which contradicts the actual findings of predictive validity (Fonteyne et al., 2021, 2019; Vanderoost et al., 2015). This misconception is an example of students' low assessment literacy (Smith et al., 2013). Students also mentioned a self-defeating prophecy effect of positioning test participation as a reason to have low trust in its predictive validity. Interestingly, no one mentioned the equally conceivable criterion contamination of a self-fulfilling prophecy (Jussim, 2017) as a reason to have more trust in the predictive validity of the test (e.g. high scorers would be more motivated and low scorers more demotivated).

Implications for Practice

Some students thought that the questions on the positioning tests were too easy, compared to the difficulty level of their study program. Some report the positioning test and associated positive feedback gave them a sense of overconfidence. This could be due to the 10 basic mathematics questions that were newly included in the 2020-2021 editions of the tests. The rationale behind introducing these basic mathematics questions is to be able to better identify at-risk students, since their predictive validity has been shown to be even greater than that of the advanced mathematics questions (Hanssens et al., 2021a). Thus, there appears to be a trade-off between predictive validity and an accurate image of the level of the study program. The latter is an important consideration. Although the feedback mentioned that a high score on the basic mathematics questions should not be viewed as a sufficient, but rather a necessary condition for academic achievement, clearer communication in this regard might be required. At the very least, a clear separation between basic and advanced mathematics questions on the test itself is recommended. Overconfidence leading to lower academic achievement is also one of the previously discussed self-defeating prophecy effects students used to explain their low trust in predictive validity.

Other aspects of the positioning test that might benefit from clearer communication to students, are the expectation of preparation, the objective of the test and its predictive validity. Current official communication does not specify if students are expected to prepare for positioning tests or not. On the other hand, questions from previous test editions are provided on the website for practice (ijkkingstoets.be). Furthermore, it did not seem clear to most

students why they should participate in the test. In order to increase motivation in such low stakes tests, according to Cole et al. (2008), communication should focus on utility and importance of participation. Communication about the predictive value of the tests could perhaps be altered in order to increase students' trust in it. Although, in doing so, one should be aware not to suppress students' growth mindset (Yeager et al., 2019), as there appears to be a previously discussed contradiction between belief in predictive validity and growth mindset among the interviewed students.

Limitations and Future Work

This study provides a rich understanding of the multitude of perceptions students have about the positioning tests. However, the following considerations limit its applicability to the general student population, other positioning tests and different (high or low stakes) assessment contexts. Firstly, there is the small sample size of the interviews. Secondly, a selection bias for the interviews can be identified on two levels: (1) only students who decided to enroll in the relevant study program were interviewed and (2) participation in the interviews was voluntary. Thirdly, it can be assumed that there are considerable differences in lived experience between participants in official and additional sessions of the positioning tests. Finally, the tests, interviews and questionnaires were taken during the exceptional time of the Covid-19 pandemic.

Interesting future work could consist of investigating perceptions of other stakeholders than test participants (e.g. test designers, SE and HE teachers), and of participants who decided not to enroll in the study program. In terms of validity research, we endorse the view of Kane (2013) that more attention should be given to consequential validity (e.g. Does face validity increase feedback adherence?) rather than criterion based validity, although the latter should not be neglected. Lastly, it would be interesting to assess in a quantitative way the various notions this study revealed (e.g. the link between context questions and perception).

Acknowledgements

This work was supported by the Flemish Interuniversity Council (VLIR) and the University of Leuven Science, Engineering and Technology Group.

References

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brown, G. T. L., & Hirschfeld, G. H. F. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy and Practice*, 15(1), 3–17. <https://doi.org/10.1080/09695940701876003>
- Cole, J. S., Bergin, D. A., & Whittaker, T. A. (2008). Predicting student achievement for low stakes tests with effort and task value. *Contemporary Educational Psychology*, 33(4), 609–624. <https://doi.org/10.1016/j.cedpsych.2007.10.002>

- Cook, D. A., Brydges, R., Ginsburg, S., & Hatala, R. (2015). A contemporary approach to validity arguments: A practical guide to Kane's framework. *Medical Education*, 49(6), 560–575. <https://doi.org/10.1111/medu.12678>
- Eccles, J. S., & Wigfield, A. (2002). Motivational Beliefs, Values, and Goals. *Annual Review of Psychology*, 53, 109-132.
- Eklöf, H. (2010). Skill and will: Test-taking motivation and assessment quality. *Assessment in Education: Principles, Policy and Practice*, 17(4), 345–356. <https://doi.org/10.1080/0969594X.2010.516569>
- Eklöf, H., & Nyroos, M. (2013). Pupil perceptions of national tests in science: Perceived importance, invested effort, and test anxiety. *European Journal of Psychology of Education*, 28(2), 497–510. <https://doi.org/10.1007/s10212-012-0125-6>
- Elwood, J. (2012). Qualifications, examinations and assessment: Views and perspectives of students in the 14-19 phase on policy and practice. *Cambridge Journal of Education*, 42(4), 497–512. <https://doi.org/10.1080/0305764X.2012.733347>
- Elwood, J., & Lundy, L. (2010). Revisioning assessment through a children's rights approach: Implications for policy, process and practice. *Research Papers in Education*, 25(3), 335–353. <https://doi.org/10.1080/02671522.2010.498150>
- Eum, K., & Rice, K. G. (2011). Test anxiety, perfectionism, goal orientation, and academic performance. *Anxiety, Stress and Coping*, 24(2), 167–178. <https://doi.org/10.1080/10615806.2010.488723>
- Finn, B. (2015). Measuring Motivation in Low-Stakes Assessments. *ETS research report series* ISSN 2330-8516, 1-17. doi:10.1002/ets2.12067
- Fitchett, S., King, K., & Champion, J. (2011). Outcomes of Mathematics Placement: An Analysis of Advising and Enrollment data. *PRIMUS-Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 21(7), 577–591.
- Fonteyne, L., Marconato, A., Tambuyzer, B., De Laet, T. De, Adriaens, M., & Melis, I. (2021). *Validiteitsrapport ijkingsstoetsen 2020*.
- Fonteyne, L., Obbels, W., Dorissen, T., Tambuyzer, B., De Laet, T. De, Adriaens, M., & Melis, I. (2019). *Validiteitsrapport Ijkingstoetsen 2019*.
- Hanssens, J., Langie, G., Van Soom, C. (2021a). Too easy? The incremental predictive validity of basic math questions on a positioning test for prospective university STEM students. Presented at the *European First Year Experience (EFYE)*, online, 07 Jun 2021-09 Jun 2021.
- Hanssens, J., Van Soom, C., Langie, G. (2021b). The predictive validity of the Engineering Technology positioning test. Presented at the *Junior Researchers of EARLI (JURE)*, online, 18 Aug 2021-20 Aug 2021.
- Hanssens, J., Van Soom, C., De Cock, M., Langie, G. (2022). Does context matter? The effect of contextualized problems on a low stakes mathematics test. Presented at *EARLI SIG1&4 Joint Conference*, Cádiz, 27 Jun 2022 - 30 Jun 2022
- Husbands, A., Rodgeron, M. J., Dowell, J., & Patterson, F. (2015). Evaluating the validity of an integrity-based situational judgement test for medical school admissions. *BMC Medical Education*, 15(1), 1–9. <https://doi.org/10.1186/s12909-015-0424-0>
- Germeijs, V., & Verschuere, K. (2006). High school students' career decision-making process: A longitudinal study of one choice. *Journal of Vocational Behavior*, 68(2), 189–204.

<https://doi.org/10.1016/j.jvb.2005.08.004>

- Jussim, L. (2017). Précis of Social Perception and Social Reality: Why accuracy dominates bias and self-fulfilling prophecy. *Behavioral and Brain Sciences*, *40*, 2013–2014. <https://doi.org/10.1017/S0140525X1500062X>
- Kane, M. T. (2013). Validating the Interpretations and Uses of Test Scores, *Journal of Educational Measurement*, *50*(1), 1–73.
- Knekta, E. (2017). Are all Pupils Equally Motivated to do Their Best on all Tests? Differences in Reported Test-Taking Motivation within and between Tests with Different Stakes. *Scandinavian Journal of Educational Research*, *61*(1), 95–111. <https://doi.org/10.1080/00313831.2015.1119723>
- Knekta, E., & Sundström, A. (2019). ‘It was, perhaps, the most important one’ students’ perceptions of national tests in terms of test-taking motivation. *Assessment in Education: Principles, Policy and Practice*, *26*(2), 202–221. <https://doi.org/10.1080/0969594X.2017.1323725>
- Liu, O. L., Bridgeman, B., & Adler, R. M. (2012). Measuring Learning Outcomes in Higher Education: Motivation Matters. *Educational Researcher*, *41*(9), 352–362. <https://doi.org/10.3102/0013189X12459679>
- Messick, S. (1995). Messick 1995 Validity of Psychological Assessment. *American Psychologist*, *50*(9), 741–749. <http://psycnet.apa.org/journals/amp/50/9/741.pdf&uid=1996-10004-001&db=PA>
- Michaelides, M. P. (2014). Validity considerations ensuing from examinees’ perceptions about high-stakes national examinations in Cyprus. *Assessment in Education: Principles, Policy and Practice*, *21*(4), 427–441. <https://doi.org/10.1080/0969594X.2014.916655>
- Nevo, B. (1985). Face Validity Revisited. *Journal of Educational Measurement*, *22*(4), 287–293.
- Pinxten, M., Van Soom, C., Peeters, C., De Laet, T., & Langie, G. (2019). At-risk at the gate: prediction of study success of first-year science and engineering students in an open-admission university in Flanders—any incremental validity of study strategies? *European Journal of Psychology of Education*, *34*(1), 45–66. <https://doi.org/10.1007/s10212-017-0361-x>
- Reich, A., Rooney, D., Gardner, A., Willey, K., Boud, D., & Fitzgerald, T. (2015). Engineers’ professional learning: a practice-theory perspective. *European Journal of Engineering Education*, *40*(4), 366–379. <https://doi.org/10.1080/03043797.2014.967181>
- Schmitt, N., Keeney, J., Oswald, F. L., Pleskac, T. J., Billington, A. Q., Sinha, R., & Zorzie, M. (2009). Prediction of 4-Year College Student Performance Using Cognitive and Noncognitive Predictors and the Impact on Demographic Status of Admitted Students. *Journal of Applied Psychology*, *94*(6), 1479–1497. <https://doi.org/10.1037/a0016810>
- Simzar, R. M., Martinez, M., Rutherford, T., Domina, T., & Conley, A. M. M. (2015). Raising the stakes: How students’ motivation for mathematics associates with high- and low-stakes test achievement. *Learning and Individual Differences*, *39*, 49–63. <https://doi.org/10.1016/j.lindif.2015.03.002>
- Smith, C. D., Worsfold, K., Davies, L., Fisher, R., & McPhail, R. (2013). Assessment literacy and student learning: The case for explicitly developing students “assessment literacy.” *Assessment and Evaluation in Higher Education*, *38*(1), 44–60. <https://doi.org/10.1080/02602938.2011.598636>
- Smither, J. W., Reilly, R. R., Millsap, R. E., Pearlman, K., & Stoffey, R. W. (1993). APPLICANT REACTIONS TO SELECTION PROCEDURES. *Personnel Psychology*, *46*, 49–76.
- Van den Broeck, L., De Laet, T., Lacante, M., Pinxten, M., Van Soom, C., & Langie, G. (2019). Predicting the

- academic achievement of students bridging to engineering: the role of academic background variables and diagnostic testing. *Journal of Further and Higher Education*, 43(7), 989–1007. <https://doi.org/10.1080/0309877X.2018.1431209>
- Vanderoost, J., Van Soom, C., Langie, G., Van Den Bossche, J., Callens, R., Vandewalle, J., & De Laet, T. (2015). Engineering and science positioning tests in Flanders: Powerful predictors for study success? *Proceedings of the 43rd SEFI Annual Conference 2015*.
- Vanderoost, Jef, Callens, R., Vandewalle, J., & De Laet, T. (2014). Engineering positioning test in Flanders : a powerful predictor for study success ? *Proceedings of the 42nd Annual SEFI Conference*, 1–8.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>
- Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., ... Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, 573(7774), 364–369. <https://doi.org/10.1038/s41586-019-1466-y>

Author Information

Jolan Hanssens

 <http://orcid.org/0000-0002-9282-9451>
KU Leuven, Leuven Engineering and Science
Education Center (LESEC), Engineering Technology
Education Research (ETHER), Faculty of Science,
Faculty of Engineering Technology
B-3000 Leuven
Belgium
Contact e-mail: jolan.hanssens@kuleuven.be

Greet Langie

 <http://orcid.org/0000-0002-9061-6727>
KU Leuven, Leuven Engineering and Science
Education Center (LESEC), Engineering Technology
Education Research (ETHER), Faculty of
Engineering Technology
B-3000 Leuven
Belgium

Carolien Van Soom

 <http://orcid.org/0000-0001-7677-0931>
KU Leuven, Leuven Engineering and Science
Education Center (LESEC), Faculty of Science
B-3000 Leuven
Belgium
