

Signal detection theory as a tool for successful student selection

Linda van Ooijen-van der Linden, Maarten J. van der Smagt, Liesbeth Woertman & Susan F. te Pas

To cite this article: Linda van Ooijen-van der Linden, Maarten J. van der Smagt, Liesbeth Woertman & Susan F. te Pas (2017) Signal detection theory as a tool for successful student selection, *Assessment & Evaluation in Higher Education*, 42:8, 1193-1207, DOI: 10.1080/02602938.2016.1241860

To link to this article: <http://dx.doi.org/10.1080/02602938.2016.1241860>



© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 19 Oct 2016.



Submit your article to this journal [↗](#)



Article views: 690



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

Signal detection theory as a tool for successful student selection

Linda van Ooijen-van der Linden^a, Maarten J. van der Smagt^b, Liesbeth Woertman^c and Susan F. te Pas^d

^aFaculty of Social and Behavioural Sciences, Department of Psychology, Departments of Experimental and Clinical Psychology, Utrecht University, Utrecht, The Netherlands; ^bFaculty of Social and Behavioural Sciences, Department of Psychology, Department of Experimental Psychology, Utrecht University, Utrecht, The Netherlands; ^cFaculty of Social and Behavioural Sciences, Department of Psychology, Department of Clinical Psychology, Utrecht University, Utrecht, The Netherlands; ^dFaculty of Social and Behavioural Sciences, Department of Psychology, Department of Experimental Psychology, Utrecht University, Utrecht, The Netherlands

ABSTRACT


Prediction accuracy of academic achievement for admission purposes requires adequate *sensitivity* and *specificity* of admission tools, yet the available information on the validity and predictive power of admission tools is largely based on studies using correlational and regression statistics. The goal of this study was to explore signal detection theory as a tool to extend the available information; signal detection theory allows for comparisons of selection outcomes on both group and individual levels and the development of tailor-made criteria for specific programmes and admission goals. We investigated who would or would not have been admitted applying specific criteria for each of three common admission tools, how many admitted students would fail and how many applicants who would have been successful would be rejected. Both comparisons at an individual level and the receiver operating characteristic curves at a group level revealed that scores obtained in a programme-specific matching programme and non-cognitive factors appear more valuable than regression statistics suggest when it comes to admitting applicants who *will* become successful students. Signal detection theory allows not only for admission-goal-specific and programme-specific fine-tuning of the content of admission tools, it also informs about the effects of criteria and thus allows the *setting* of criteria.

KEYWORDS

Admission; student selection; signal detection theory; criterion; non-cognitive factors

Introduction

Accurate predictions of academic achievement informing admission decisions require valid admission tools. The validity and predictive power of admission tools have received much interest over the years (Bingham 1917; Richardson, Abraham, and Bond 2012; Kotzee and Martin 2013), but the sensitivity and specificity of the presently available tools is not satisfactory. Secondary school Grade Point Average (SSGPA), a measure of past cognitive achievement, is the most commonly used indicator of future academic achievement in selection and admission procedures. In addition to SSGPA, universities in North America use standardised admission instruments such as SAT (the Scholastic Aptitude Test) and ACT (American College Testing, these days broadly supporting education and workforce development). In Europe, university programmes and secondary education systems, as well as their admission procedures,

CONTACT Linda van Ooijen-van der Linden  l.vanoijen@uu.nl

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

differ across countries. In the Netherlands, for instance, secondary education is divided into three distinct levels, of which only the highest gives direct access to university (De Witte and Cabus 2013).

Admission to university is therefore largely determined by *past* academic achievement. Given that it is often considered the best available predictor of *future* academic achievement, this is logical (Richardson, Abraham, and Bond 2012), yet ‘the best available’ in university admissions does not equal a prediction of academic achievement that can be called accurate (Sternberg et al. 2012; Sawyer 2013). Although Shulruf, Hattie, and Tumen (2008) (see also Shulruf et al. 2011) found correlations >0.60 of SSGPA with university first-year GPA using a specific and new calculation of SSGPA, it is more commonly reported as only a reasonable predictor of university GPA, with $r \approx 0.40$ (Richardson, Abraham, and Bond 2012; Edwards, Friedman, and Pearce 2013). Although their predictive power in most studies is arguably much smaller than desired, past academic achievement (De Koning et al. 2012; Van der Heijden, Hessen, and Wubbels 2012), cognitive admission tests (Visser et al. 2012) and non-cognitive measures (Huang 2012; Richardson, Abraham, and Bond 2012; Sternberg et al. 2012) have all been shown to be valid predictors of academic achievement to a certain extent (Robbins et al. 2004; O’Neill et al. 2011; Sinha et al. 2011; Poole et al. 2012; Schmitt 2012; Stemler 2012; Kreiter and Axelson 2013; Preiss et al. 2013; Sawyer 2013; Christersson et al. 2014; Sartania et al. 2014; Simpson et al. 2014; Gaertner and Larsen McClarty 2015).

Thus, the science establishing evidence-based admission tools for university programmes is still mainly a work in progress. Admission procedures do appear to be converging on the inclusion of multiple cognitive and non-cognitive admission scores (including optional questions applicants can choose to answer to show their academic potential) as ‘best available practices’ (Cortes 2013). Knowing which predictors of academic achievement are valid, however, does not answer the question of which specific criteria to use for each or all predictors. Whether highly selective or more liberal criteria are appropriate, differs across programmes and the goals set by admission committees.

We explored applying signal detection theory to a cohort of psychology freshmen and a newly developed matching programme. Signal detection theory (Green and Swets 1966; Stanislaw and Todorov 1999; Macmillan and Creelman 2005) might help to investigate the sensitivity and specificity of a set of admission instruments. The main reason to explore this application is the pay-off matrix it provides, which can help admission committees to make informed decisions on the cut-off value or criterion. Over and above the information regression analyses provide, the pay-off matrix offers information on the results of decisions, the proportions of both successful and unsuccessful students that would have been admitted and the proportions of both successful and unsuccessful students that would have been rejected applying a certain criterion. In addition, signal detection theory allows analyses at an individual level.

A valid predictor as such does not prescribe whether a score of 6.4 or 7.0 should be a more useful cut-off criterion than any other value for the decision on admission. Admission committees set goals and choose a criterion to match these goals (Sternberg et al. 2012; Cortes 2013; Sawyer 2013), and signal detection theory offers a way to translate goals into evidence-based criteria. Preferably most admitted students will do fine (in signal detection theory terminology these are deemed ‘hits’). However, some will inadvertently fail (‘false alarms’). Of the rejected applicants some would indeed have failed had they been admitted (these are considered ‘correct rejections’), while others would have done fine (‘misses’). The relative importance of maximising hits or correct rejections, and of minimising false alarms or misses, can be pre-determined by admission committees in a pay-off matrix (see Table 1), which then prescribes an admission criterion, called c in the simplest

Table 1. The pay-off matrix applied to admission decisions.

| | | Decision | | Total (%) | |
|----------------------|--------------|-------------|-------------------|-----------|-------------|
| | | Admit | Reject | | |
| Academic achievement | Successful | Hit | Miss | 100 | Sensitivity |
| | Unsuccessful | False alarm | Correct rejection | 100 | Specificity |

signal detection theory model, given the sensitivity and specificity of the admission tools. At zero the criterion reflects neutrality, that is no preference in minimising misses or false alarms and thus equal proportions of both.

A criterion chosen to prevent sending away students who would probably have been successful is less selective, resulting in an admission-bias, which is reflected in $c < 0$. If the goal is to prevent admission of students who will likely fail, a more selective criterion is chosen, resulting in a rejection-bias: $c > 0$. This is a prescriptive use of the criterion; one decides on desirable relative weights of misses and false alarms in the pay-off matrix before making the admission decisions and setting the criterion to match this decision. The criterion can also be used descriptively; after making a number of decisions, in this case on hypothetically admitting matched applicants, the criterion used results in a pay-off matrix, showing if there is a bias and if so, what type and size.

In signal detection theory, both hit-rate and miss-rate, as well as false-alarm-rate and correct-rejection-rate sum up to 1. In the simplest model, using the hit-rate (successful admitted students) and false-alarm-rate (unsuccessful admitted students), d' can be calculated. Although often being referred to as a measure of sensitivity, d' actually combines sensitivity (hits/(hits + misses)) and specificity (correct rejections/(correct rejections + false alarms)), and indicates how well one can discern successful students from unsuccessful students. In this context, d' is the (standardised) distance between the mean score on (a set of) admission tests of successful students and the (standardised) mean score of unsuccessful students, assuming the variance of the distribution of the scores of successful students is equal to that of unsuccessful students.

Plotting the hit-rates against the false-alarm-rates in a receiver operating characteristic curve visualises the accuracy, which can range from 0.5, chance level, to 1, perfect accuracy. A' is an *estimation* of the area under the receiver operating characteristic curve based on only one hit-rate and false-alarm rate (Macmillan and Creelman 2005). A' can also vary from 0.5, chance level, to 1, perfect discernibility. In the case of admission test scores that are not normally distributed a non-parametric measure of sensitivity, such as A' , is more suitable. B'' is a non-parametric criterion which is also calculated from the hit-rate and the false-alarm-rate. B'' ranges from -1 to 1 , with $B'' = 0$ reflecting a lack of bias, values < 0 reflecting an admission-bias and values > 0 reflecting a rejection-bias.

The advantage of using signal detection theory over regression is that the latter can only tell which predictors are useful, and subsequently rank them according to their degree of usefulness for the given programme. Through the pay-off matrices signal detection theory enables fine-tuning of the admission criterion based on the relative weight of misses and false alarms; the descriptive matrices for each possible criterion of the previous cohorts can be compared, enabling prescriptive use for a future cohort. In addition, it allows for analyses at the individual level, since every student obtains an admission score that can be compared with both a (set) criterion and the scores of other students.

In the present study, data from a special cohort of applicants were studied (tentatively starting in 2013) where Utrecht University offered psychology applicants a non-obligatory matching programme. This offered the opportunity to investigate not only the sensitivity but also the specificity of the admission tools used. The purpose of the present study is to explore the application of signal detection theory in university admission to a psychology programme, with regard to making informed decisions on the content of the admission tools and what criteria to use.

Methods

Participants

In September 2013, 524 students started their freshman year in psychology at Utrecht University (UU), The Netherlands. Of them 288 had participated in a four-day matching programme in April and May of that year (matched students), whereas 236 had not (non-matched students). Of the matched students 233 were female with a mean age of 19.2 years ($SD = 1.2$), 55 were male with a mean age of 20.1 years ($SD = 2.0$). Of the students who did not participate in the matching programme 172 were female with

a mean age of 20.5 years ($SD = 4.2$), 64 were male with a mean age of 20.2 years ($SD = 1.7$). Students were encouraged to participate in the matching programme, but it was not obligatory. The matching programme in 2013 served as a pilot before it became obligatory and part of the admission procedure in 2014.

Matching programme

The matching programme was developed to be representative of the courses in both quartile 1 (Q1; the semester at UU is split into two blocks of ten weeks, so the academic year has two semesters and four quartiles) of the freshman year and the freshman year as a whole. During the first of four consecutive days, applicants followed two lectures on two main topics and a tutorial on both. The middle two days were for home study. On the last day, the students submitted an essay and completed a multiple-choice test consisting of 20 questions, 10 about each main topic. In June, all students received a letter with their examination and essay grades.

Measures

Predictors

The choice to include achievement at secondary school, cognitive admission test scores and non-cognitive measures as predictors of academic achievement was based on the literature (e.g. Richardson, Abraham, and Bond 2012; Cortes 2013). However, the matching programme and the university wide application form were developed prior to the idea of investigating the application of signal detection theory and publishing. The application form included the grades for the four obligatory secondary school topics in the Netherlands – mathematics, Dutch, English and social sciences – plus biology given its relevance for psychology. The matching tests were designed to represent the type and difficulty of

Table 2. Overview of scoring of the predictors in the admission toolkits and the participants' scores.

| Admission toolkit | Scoring | Matched students | Non-matched students |
|--|---|----------------------|----------------------|
| | | Mean (SD) – <i>n</i> | Mean (SD) – <i>n</i> |
| SSGPA | Average of Mathematics, English, Dutch, Biology and Social Sciences | 6.8 (0.5) – 279 | 6.8 (0.5) – 178 |
| Matching scores (MS) | | 6.1 (1.8) – 288 | |
| Essay | Ten-point scale | 6.2 (1.4) – 306 | |
| MC questions Topic 1 | # correct answers, max 10 | 7.3 (1.7) – 392 | |
| MC questions Topic 2 | # correct answers, max 10 | 6.0 (1.8) – 392 | |
| Non-cognitive factors (NCF) | Ten-point scale | 5.7 (2.1) – 285 | 4.4 (2.8) – 188 |
| Grade goal | 0–4 | 3.0 (0.8) – 285 | 2.4 (1.4) – 188 |
| 'I prefer not to aim for a C, but rather work for a better grade' | 0 = disagree strongly, 4 = agree strongly | | |
| Extra examination course | 1 if done, 0 if not | 142 (of 288) | 86 (of 236) |
| 'Did you take an extra examination course? (in addition to the ones required for your certificate)' | | | |
| Interest in honours programmes | 1 if 'yes', 0 if 'not' or 'maybe' | 46 (of 288) | 34 (of 236) |
| 'Are you interested in the honours programme of your bachelor's programme?' | | | |
| Community commitment" 'Were you involved in extra activities? (e.g. student council, school magazine, community work)' | 1 if 'yes' according to the open question about extra activities | 80 (of 288) | 50 (of 236) |

Note that several dozens of applicants did not submit an essay and other students did submit an essay but did not take the examination. Only students who did both are considered 'matched' and therefore the numbers differ. Nine of the matched students did not fully complete the matching form. Of the 236 non-matched students 178 partially filled in the matching form. An additional 10 completed the form.

Table 3. Description of four admission toolkits; the three single predictors, SSGPA, MS and NCF and their unweighted averages. All range 0–10.

| AT | Unweighted averages |
|-----------------------------|--|
| SSGPA | (Mathematics + English + Dutch + Biology + Social Sciences)/5 |
| Matching average (MS) | (Matching assignment + Examination Topic 1 + Examination Topic 2)/3 |
| Non-cognitive factors (NCF) | (Grade goal + community commitment + extra examination course + interest in honours) $\times 1.43$ |
| SSGPA/MS/NCF | (SSGPA + MS + NCF)/3 |

tests in the freshman year – a multiple-choice examination and an individually written essay – of course with the limitations inherent to ‘studying for a week’.

From the available non-cognitive measures in the application form, we chose the ones reported in Richardson, Abraham, and Bond (2012) and Cortes (2013) to be most predictive of academic achievement: academic intrinsic motivation, grade goal and community commitment (see Tables 2 and 3). We used three types of predictors; secondary education grades (SSGPA), matching scores (MS) and non-cognitive factors (NCF). These three predictors, SSGPA, MS and NCF, will be called admission toolkits in the remainder of this article. Their combination, SSGPA/MS/NCF, is also considered an admission toolkit.

Academic achievement

In order to check the stability of the prediction of academic achievement within this cohort, academic achievement was operationalised fourfold; Q1-first – grades for first attempts of assignments and examinations in quartile 1, Q1-final – whether both courses in quartile 1 were successfully completed after resits, S1-final – whether all four (obligatory) courses in the first semester were successfully completed after resits and Y1 – whether at least seven of the eight courses in year 1 were successfully completed.

Analyses

Correlations, regression analyses and odds ratios served as a comparison ground for the signal detection theory outcomes. Non-parametric analyses were run because the distributions of several scores obtained in the matching programme were skewed. Spearman's rho correlations were calculated to compare the predictive validity of our admission toolkits with previous findings. A backward logistic regression with Y1 as the dependent variable was run to test whether SSGPA, MS and NCF would all independently explain variance in academic achievement. Another logistic regression was run to test how well the unweighted average of SSGPA, MS and NCF predicted Y1. Odds ratios were computed to bridge the differences in interpretation between correlation and regression and signal detection theory, and to aid the interpretation of the distribution of students over hits, false alarms, misses and correct rejections.

The three single admission toolkits and their unweighted average were tested for their sensitivity applying signal detection theory in differentiating successful students from unsuccessful students. This was done by taking each admission toolkit (see Table 3) and using these as individual admission scores for each applicant. For each of the four admission toolkits, all practically reasonable scores were used as the cut-off score to decide for each applicant whether they would be admitted using this admission toolkit and this criterion. Since all applicants were admitted and all data of their freshman year were available, we determined for each individual and for each operationalisation of academic achievement whether they should be considered a hit, a miss, a false alarm or a correct rejection for each combination of admission tool and criterion. This resulted in different numbers of hits, misses, false alarms and correct rejections for different criteria; more hits, fewer misses, more false alarms and fewer correct rejections when admitting applicants using relatively low criterion values (being less selective, $B'' < 0$), compared to fewer hits, more misses, fewer false alarms and more correct rejections when using higher criterion values (being more selective, $B'' > 0$). These outcomes were plotted for each admission toolkit for each criterion in bar charts and receiver operating characteristic curves. The

Table 4. Backward logistic regression with Y1 as the dependent variable. $n = 279$.

| | 95% CI for odds ratio | | | |
|---------------------------|-----------------------|-------|------------|-------|
| | B (SE) | Lower | Odds ratio | Upper |
| <i>Single predictors</i> | | | | |
| Step 1 | | | | |
| Constant | -7.00 (2.53) | | | |
| SSGPA | 1.05 (0.40)* | 1.31 | 2.87 | 6.23 |
| MS | 0.09 (0.09) | 0.91 | 1.09 | 1.31 |
| NCF | 0.26 (0.14) | 0.99 | 1.30 | 1.69 |
| Step 2 | | | | |
| Constant | -6.77 (2.50) | | | |
| SSGPA | 1.09 (0.40)* | 1.37 | 2.96 | 6.43 |
| NCF | 0.28 (0.14)** | 1.01 | 1.32 | 1.72 |
| <i>Unweighted average</i> | | | | |
| Constant | -2.07 (0.94) | | | |
| SSGPA/MA/NCF | 0.61 (0.16)* | 1.34 | 1.83 | 2.50 |

Notes: Single predictors: $R^2 = 0.07$ (Cox & Snell), 0.12 (Nagelkerke). Model $\chi^2(3) = 21.23$, $p < 0.001$.

* $p < 0.01$.

** $p = 0.04$.

Unweighted average: $R^2 = 0.06$ (Cox & Snell), 0.09 (Nagelkerke). Model $\chi^2(1) = 15.85$, $p < 0.001$.

* $p < 0.001$.

area under the receiver operating characteristic curve is calculated for each admission toolkit as the most appropriate measure of accuracy. A 's are reported to aid comparison of more complex admission toolkits with specific criteria in Figure 4.

Next, to compare the outcome of admission toolkits predicting Y1 at different degrees of selectivity, for each admission toolkit the number of hits, misses, false alarms and correct rejections was plotted for the criterion resulting in admission of 70, 50 or 30% of applicants. The number of students selected *uniquely* based on SSGPA, MS or NCF for 70, 50 and 30% admission (predicting Y1), and how many of them actually were successful (hits), was calculated to assess the independent predictive value of SSGPA, MS and NCF, applying signal detection theory.

Results

Replication of findings on SSGPA, MS and NCF

The SSGPA based on the matching form (obtained before admission to the university) correlates with the final and official SSGPA, $r_s(254) = 0.77$, $p < 0.001$, and final SSGPA also correlates with the number of courses completed in Y1, $r_s(256) = 0.41$, $p < 0.001$. The SSGPA based on the matching form correlates with the number of courses completed in year 1 (NoCC-Y1) as well, $r_s(279) = 0.32$, $p < 0.001$. NCF correlates with NoCC-Y1, $r_s(285) = 0.27$, $p < 0.001$. MS, a more momentary measure of academic functioning, correlates with NoCC-Y1, $r_s(288) = 0.20$, $p = 0.001$. The correlation of SSGPA/MS/NCF with NoCC-Y1 was $r_s(279) = 0.30$, $p < 0.001$. These correlations concur with earlier findings (Richardson, Abraham, and Bond 2012). A backward logistic regression starting with SSGPA, MS and NCF in the model resulted in a model with SSGPA and NCF as significant predictors of Y1, explaining 7–12% of the variance (see Table 4). The unweighted average of all three predictors appeared to be of comparable predictive value, with 6–9% of the variance explained.

Of the 288 matched students, 83% were successful (completed at least seven of eight freshman year courses) and 17% were not successful according to this operationalisation of academic achievement. Given the medium sample size and the small group of unsuccessful students, odds ratios probably are better indicators of the effect sizes than r and R^2 . For SSGPA, MS and NCF, the odds ratios were 2.87, 1.09 and 1.32, respectively. SSGPA/MS/NCF had an odds ratio of 1.83.

Table 5. Odds ratios and their calculations – based on signal detection theory outcomes – for the four admission toolkits.

| | Hits/ misses | False alarms/ correct rejections | Odds ratio |
|-----------------------|----------------|----------------------------------|------------------|
| SSGPA ≥ 7 | 111/120 = 0.93 | 9/39 = 0.23 | 0.93/0.23 = 4.01 |
| MS ≥ 5.6 | 174/57 = 3.05 | 25/23 = 1.09 | 3.05/1.09 = 2.81 |
| NCF ≥ 5.7 | 147/84 = 1.75 | 15/33 = 0.45 | 1.75/0.45 = 3.85 |
| SSGPA/MS/NCF ≥ 6 | 147/84 = 1.75 | 16/32 = 0.50 | 1.75/0.50 = 3.50 |

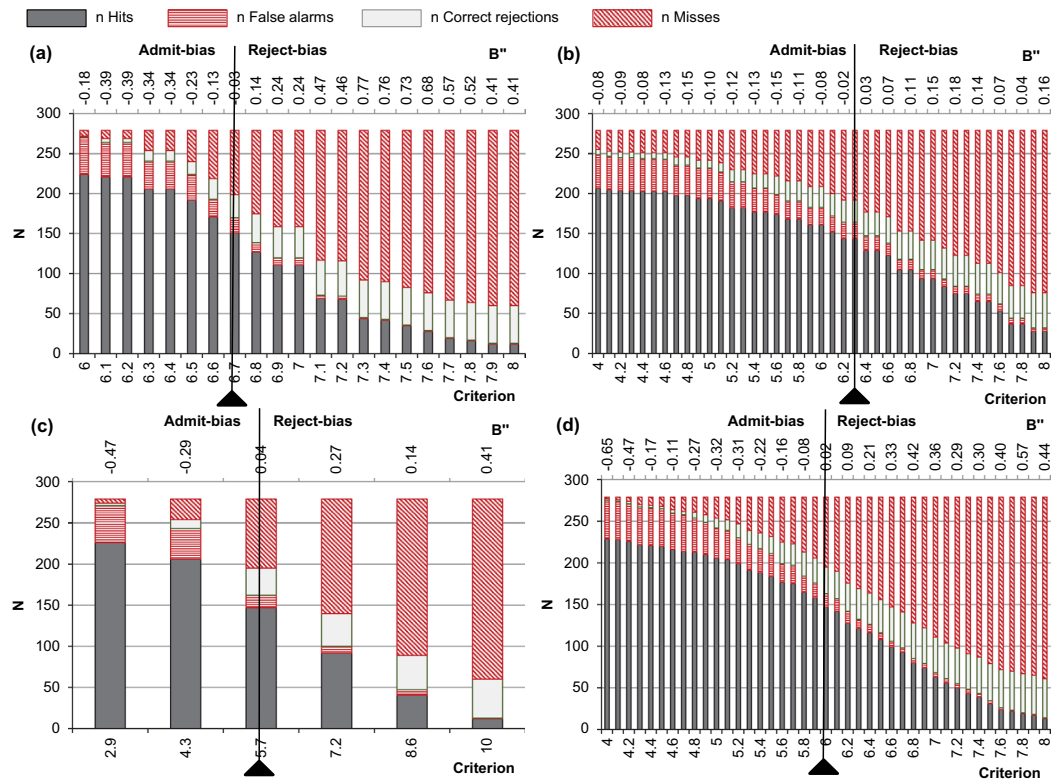


Figure 1. Signal detection theory outcomes for all applicants who participated in the matching programme ($n = 279$) for the four admission toolkits (panel a–d, see also Table 3). The admission scores, which are the criteria on the bottom x-axis, are predicting Y1. Students with a score equal to or higher than the value on the lower x-axis were hypothetically selected. For each score/criterion, the y-axis gives the total number of students and their distribution over hits, false alarms, correct rejections and misses. a – SSGPA, b – MS c – NCF, d – SSGPA/MS/NCF.

Signal detection theory outcomes at group level

To check for the stability of the selection outcomes, all four operationalisations of academic achievement (Q1-first, Q1-final, S1-final and Y1) were plotted. All three single predictors showed stability of the selection outcomes over the year, and thus we report results for Y1 only. The signal detection theory approach replicated both the validity of SSGPA, MS and NCF (see Tables 4 and 5, Figures 1 and 2) as single predictors. Figures 1 and 2 show that the combined admission toolkit consisting of all three predictors did not lead to any major improvements in terms of sensitivity and specificity, or in terms of differential effects of different criteria. The area under the curve did not increase when combining predictors, but the current study did replicate earlier findings on SSGPA, MS and NCF showing them to be valid predictors of Y1 (Richardson, Abraham, and Bond 2012). It also replicated the finding that adding predictors does not result in major improvements of predictive power. Odds ratios based on the numbers of hits, misses, false alarms and correct rejections are given in Table 5.

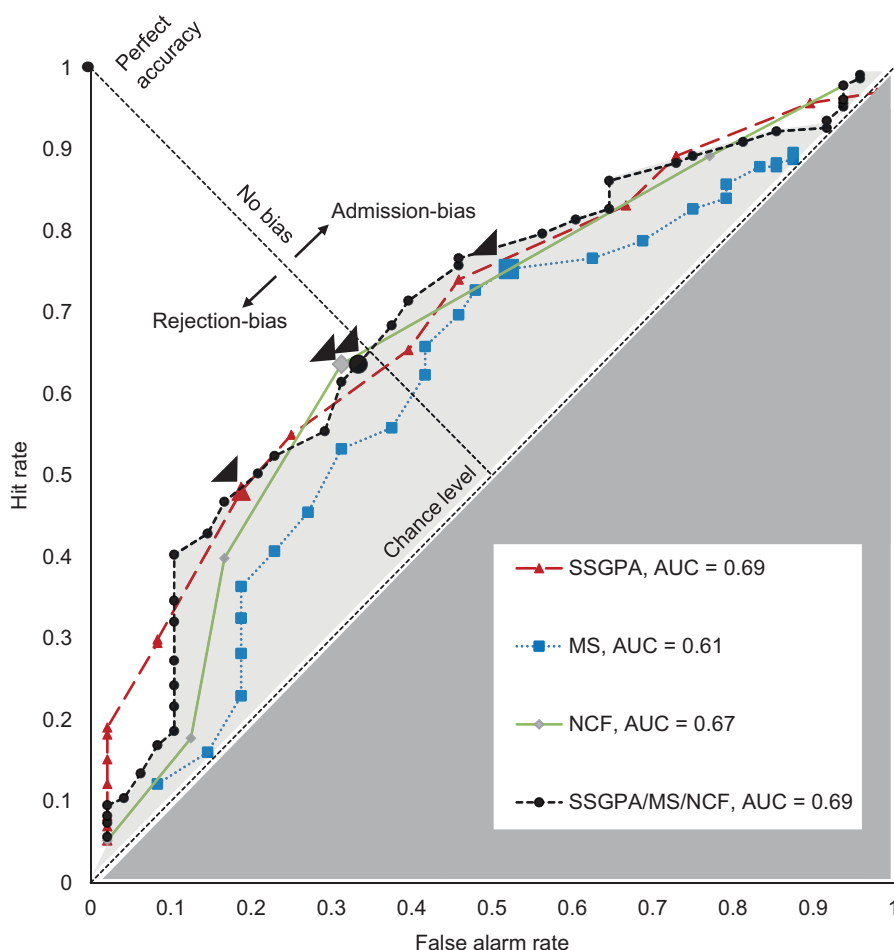


Figure 2. Receiver operating characteristic curves for admission toolkits SSGPA, MS, NCF and SSGPA/MS/NCF ($n = 279$). In these receiver operating characteristic curves the Hit-rate is plotted against the False-Alarm-rate for each criterion. At the diagonal from (0, 0) to (1, 1), the admission toolkit performs at chance level in discerning successful students from unsuccessful students. At the upper left corner where the False-Alarm-rate is zero and the Hit-rate 1, the accuracy of the admission toolkit is perfect. The area below and right from the diagonal is the receiver operating characteristic-space of accuracy below chance level and therefore irrelevant. The area above and left from the diagonal is the receiver operating characteristic-space of accuracy above chance level. The grey parts visualise the area under the curve (AUC) of 0.69 of SSGPA/MS/NCF. The other straight diagonal line, from (0, 1) to the middle, reflects neutral criteria; points above (and further away) from this line reflect a (stronger) admission bias, points below (and further away) from this line reflect a (stronger) rejection bias.

The odds of being admitted were 4.01 times higher for successful students than for unsuccessful students for $\text{SSGPA} \geq 7$, and 2.81 times higher for successful students than for unsuccessful students for $\text{MS} \geq 5.6$. For $\text{NCF} \geq 5.7$ successful students were 3.85 times more likely to be admitted than unsuccessful students, and for $\text{SSGPA/MS/NCF} \geq 6$ the odds ratio was 3.50. Following the recommendations of Ferguson (2009), the odds ratios of all three predictors and their unweighted average would represent moderate to strong effects. For MS, NCF and SSGPA/MS/NCF, the odds ratios based on the signal detection theory outcomes – 2.81, 3.85 and 3.50 – were larger than the upper limits of 95% confidence intervals of the odds ratios based on the regression analysis; 1.31, 1.72 and 2.50. Please note that $(\text{SSGPA} + \text{MS} + \text{NCF})/3$ has a wider range of possible scores than SSGPA, since lower scores are both possible and common for MS and NCF, whereas they are not for SSGPA (Table 2). This is due to, in The Netherlands, one being unable to enter the examination year or finish school with $\text{GPA} < 6$. Only one five is allowed for Dutch, English or mathematics. Thus, the range of SSGPA, even for the pre-final year,

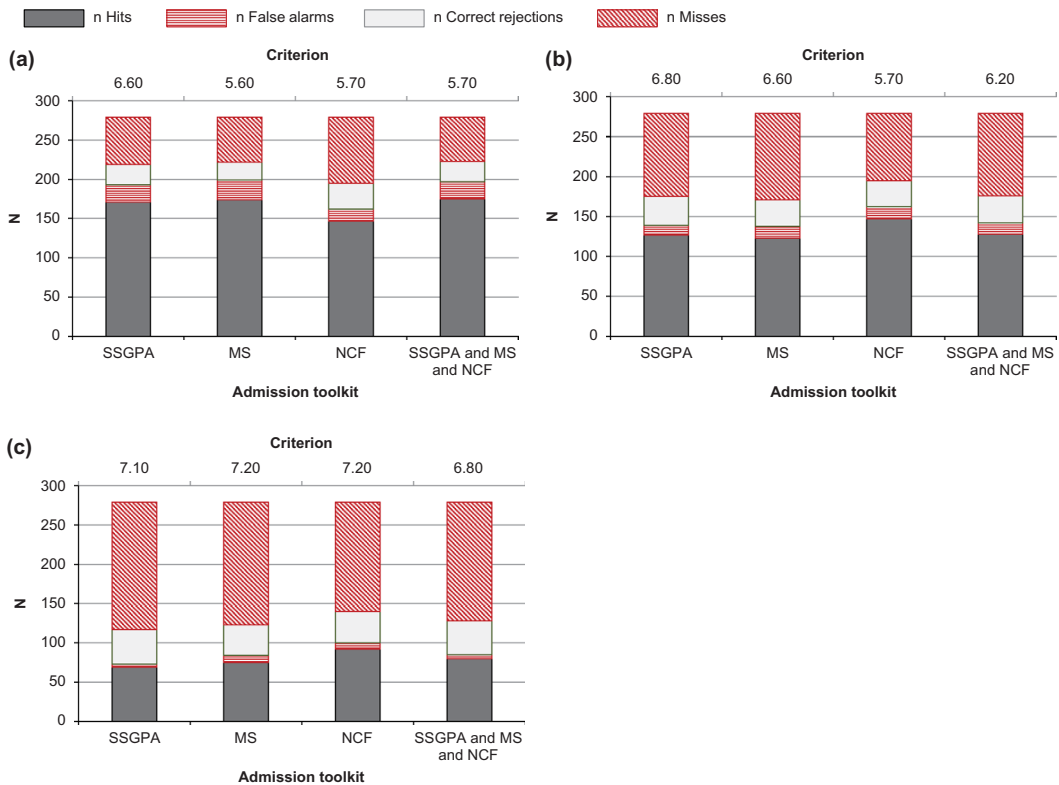


Figure 3. A comparison of the characteristics of the four admission toolkits on admission of 70, 50 or 30% of all applicants who participated in the matching programme, all predicting Y1 ($n = 279$). The bottom x-axis states each admission toolkit that was put in signal detection theory; students with a score equal to or higher than the value on the upper x-axis were hypothetically selected, students with a lower score were hypothetically rejected. For each score/criterion, the y-axis gives the total number of students and their distribution over hits, false alarms, correct rejections and misses. a – 70% admission, b – 50% admission, c – 30% admission.

starts at six. This means that adding MS and NCF to the toolkit increases the decision space: the range of grades is wider and so is the range of possible cut-off grades.

The unweighted average of SSGPA, MS and NCF yielded an A' of 0.73 at criterion 6 which was neutral with $B'' = 0.02$. Using this criterion would result in admitting 58% of matched students, as well as 64% correct decisions. Setting the criterion to admit 70% of applicants, $((SSGPA + MS + NCF)/3) \geq 5.7$, results in outcomes that are comparable to that of the other admission toolkits ($B'' = -0.15$ and 72% correct decisions and see Figure 3 panel (a), but with the inclusion of both programme specific and non-cognitive admission criteria. In this small sample with a small distribution for SSGPA, the signal detection theory approach provided more detailed information than the correlation and regression analyses.

Using signal detection theory-outcomes to set tailor-made criteria

Since the criteria required to admit 70% of the applicants differed across admission toolkits, an exploration of the effects of combining single admission toolkits with *specific* criteria into double or triple admission toolkits (instead of SSGPA/MS/NCF) might be fruitful. Hence, for each of the single admission toolkits a criterion was chosen that appeared to best fit its primary strength in optimising admission decisions. Starting with SSGPA, the lower three bars of Figure 4 show that SSGPA with a relatively high criterion (7) was suitable for restricting the proportion of false alarms, yet at the cost of a relatively high proportion of misses. For MS a lower criterion (5.6) resulted in a relatively high proportion of hits, and

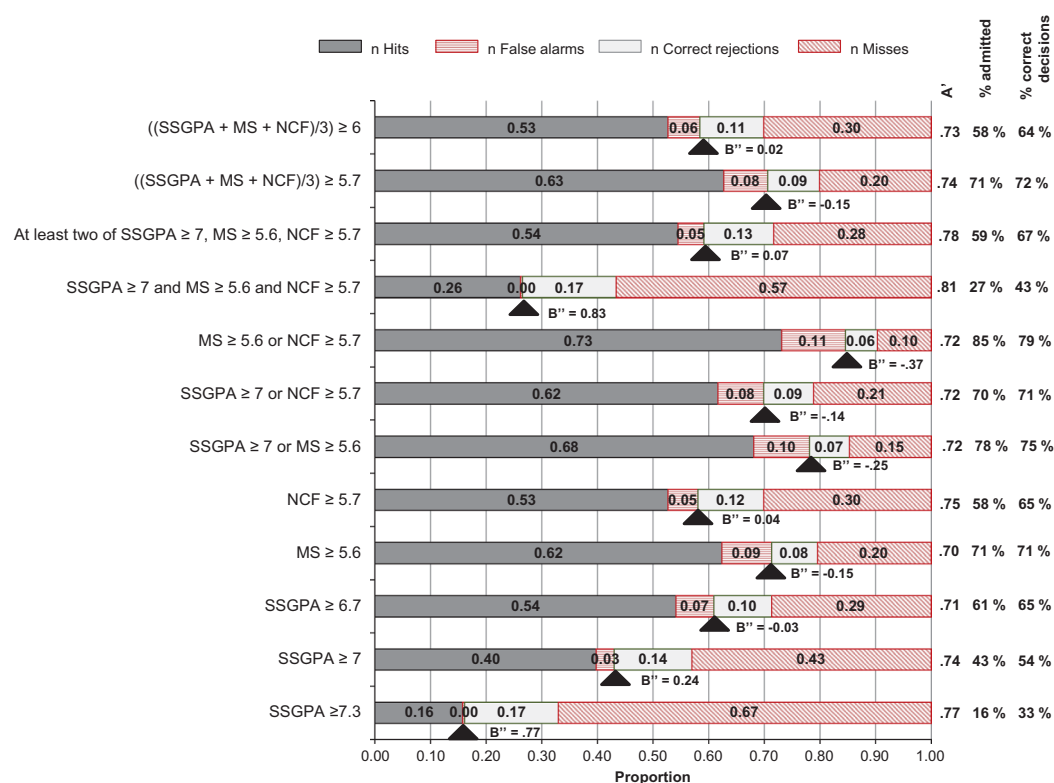


Figure 4. A comparison of admission toolkits based on the cut-offs of 7, 5.6 and 5.7 for SSGPA, MS and NCF, respectively, predicting Y1 ($n = 279$). For 'At least two' matched students were hypothetically admitted if they scored ≥ 7 , ≥ 5.6 or ≥ 5.7 on at least any two of SSGPA, MS or NCF, respectively. $((SSGPA + MS + NCF)/3) \geq 6$, $((SSGPA + MS + NCF)/3) \geq 5.7$, SSGPA ≥ 6.7 and SSGPA ≥ 7.3 were added for comparison. The criterion and its results in terms of proportion of admitted applicants, given on the x-axis, is represented by the black triangle for each admission toolkit.

at the same time a lower proportion of misses compared to SSGPA (7) (compare the four lower bars of Figure 4). Performance of NCF set at 5.7 was intermediate between SSGPA (7) and MS (5.6), both in admission percentage and the percentage correct decisions; see the fifth bar from the bottom in Figure 4.

In other words, for these admission toolkits and criteria, SSGPA showed a rejection bias, MS an admission bias, while NCF was neutral, as is also illustrated in the receiver operating characteristic curves (the enlarged data points with black triangles in Figure 2). Whereas in Figure 3 for each admission toolkit the signal detection theory-outcomes are shown for one specific criterion, the information on the proportions of hits and false alarms is given for all criteria in Figure 2, resulting in a receiver operating characteristic curve. Note that receiver operating characteristic curves provide a clear overview of the accuracy of different admission toolkits across criteria, and allow us to determine which ratio of hits and false alarms – and thus which criterion – matches the goals of the admission committee.

As apparent from Figure 4 (sixth, seventh and eighth bars from the bottom), double admission toolkits requiring a score passed the cut-off for only one of two single admission toolkits, appeared to show a slight performance improvement over single admission toolkits in percentages of admitted applicants and correct decisions. As could be expected, a combination of all three single admission toolkits led to the highest A' (hardly any false alarms), but mainly due to the low percentage admittance of SSGPA (7); only 27% of students would be admitted (57% misses, only 43% correct decisions). An admission toolkit admitting students who pass the criterion for *at least two* of the three single admission toolkits (third bar from the top) resulted in a lower proportion of false alarms and higher proportion of misses, when compared to the double admission toolkits (compare the third bar from the top with the sixth,

Table 6. Number of students that would be considered unique hits and false alarms for SSGPA, MS or NCF only, with criteria set to admit 70, 50 and 30% of matched applicants.

| Admission rate (%) | SSGPA | MS | NCF | Total |
|----------------------------|-------|----|-----|-------|
| <i>Unique hits</i> | | | | |
| 70 | 13 | 25 | 7 | 45 |
| 50 | 15 | 19 | 29 | 63 |
| 30 | 14 | 27 | 35 | 76 |
| <i>Unique false alarms</i> | | | | |
| 70 | 6 | 7 | 4 | 17 |
| 50 | 6 | 6 | 7 | 19 |
| 30 | 3 | 5 | 5 | 13 |

Note that for NCF the criterion lies at 5.7 for both 70 and 50% admission at which 58% is admitted.

seventh and eighth from the bottom). 'At least two' showed a slightly higher A' than the double admission toolkits, and slightly lower but still reasonable percentages of admittance and correct decisions. A comparison with SSGPA/MS/NCF, also reported in the rightmost bars of Figure 3, revealed 'at least two' provided no improvements over SSGPA/MS/NCF ≥ 6 : the differences in sensitivity, specificity, B'' , percentage admittance and percentage correct decisions are small to non-existent.

The neutral criteria – $B'' = 0.02$ for SSGPA/MS/NCF ≥ 6.0 , $B'' = 0.07$ for 'at least two' and $B'' = 0.04$ for NCF (5.7) – mean that these admission toolkits did not prioritise minimising one type of incorrect decision, misses or false alarms, over the other. Note that the A 's of 0.73, 0.78 and 0.75 lie at 46–56% of the range between chance level – 0.5 – and perfect predictions – 1 – in predicting academic achievement.

Signal detection theory outcomes at the individual level

Comparing admission toolkits on all signal detection theory outcomes (see Figure 3) did not provide clear-cut information aiding in the choice of predictors and criteria. Yet, selectivity was not the same for all admission toolkits, since setting the criterion to admit 70% of applicants resulted in different criteria; i.e. 6.6 for SSGPA, 5.6 for MS, 5.7 for NCF and 5.7 for the combination of all three (Figure 3).

A comparison of selection outcomes for SSGPA, MS and NCF on an individual level, as is shown in Table 6, revealed that the different predictors select slightly different subsets of individuals. When selecting 70% of applicants, 62 students passed the criterion for only SSGPA, MS or NCF when comparing these three single-predictor admission toolkits; 19 for SSGPA, 32 for MS and 11 for NCF. Forty-five of them turned out to be successful (hits); 13 for SSGPA, 25 for MS and seven for NCF. As selectivity is increased by setting the criteria to admit only 50% of matched applicants (Figure 3, panel (b)), SSGPA yielded 15 unique hits and six unique false alarms, MS yielded 19 unique hits and six unique false alarms. Hypothetically admitting 30% of applicants (Figure 3, panel (c)) resulted in 14 and 27 unique hits for SSGPA and MS (and three and five false alarms, or six and seven at 70% admission). For NCF the unique hits increased from seven to 35 with increasing selectivity from 70 to 30%. The number of false alarms increased from four to five. These data suggest it is worthwhile to use multiple indicators of future academic achievement.

Discussion

We replicated earlier findings on SSGPA, MS and NCF as predictors of academic achievement (e.g. De Koning et al. 2012; Richardson, Abraham, and Bond 2012; Visser et al. 2012), and we have explored signal detection theory as a tool to evaluate selection outcomes across different admission toolkits and criteria. The regression analyses and odds ratios appeared to indicate MS might be less valuable in admission toolkits than SSGPA and NCF. Note that the odds ratios based on signal detection theory outcomes give comparisons between successful and unsuccessful students, whereas regression analyses provide information on the group as a whole. The odds ratios based on the signal detection theory outcomes

appeared to indicate stronger effects than the correlations and regression analyses, but, taking into account that successful students outnumbered the unsuccessful students by a factor of four, leaves an odds ratio of four equivalent to around 10% explained variance, which can be meaningful, but is considered a small effect (Ferguson 2009). Thus, odds ratios provide relevant information in addition to regression analyses, but still only at the group level.

Signal detection theory outcomes

Signal detection theory helped to assess the validity of admission tests by showing stability of selection outcomes as time passes within the freshman year. A close look at sensitivity, specificity and the pay-off matrices of different admission toolkits was taken in an attempt to contribute to better informed decisions in selection. The accuracy of selection based on SSGPA, MS and NCF was not as high as would be desirable, but using all three admission toolkits did widen the range of the admission scores and increased the resolution of the decision space. Additionally, the receiver operating characteristic curves and the signal detection theory outcomes at the individual level suggested MS might after all be valuable in addition to SSGPA in predicting academic achievement.

The selectivity of the different admission toolkits differed at the same criteria. Together with the finding that several dozens of students scored above criterion on only one of the single admission toolkits when admitting 70% of applicants, this suggests that individual differences might exist in what are adequate predictors of academic achievement. A further comparison of individual outcomes across admission toolkits revealed that, partly depending on the selectivity of the chosen criterion, NCF especially might be a valuable admission toolkit in terms of minimising rejection of applicants who did not do too well in secondary education and the matching programme, but who *will* become successful students. For another subset of applicants, SSGPA and MS were the admission toolkits that predicted academic success, while the other two, MS and NCF or SSGPA and NCF, did not. The finding that each predictor at low, medium as well as high levels of selectivity *uniquely* admits applicants, most of whom *will* become successful students, suggests the 'best available' toolkit includes all three admission toolkits.

The receiver operating characteristic curves clearly showed the near equivalent accuracy of the single admission toolkits and the combined admission toolkit. Combining SSGPA with other predictors in an admission toolkit did not heighten accuracy, but it did increase decision space and allows the detection and selection of potentially successful applicants on other characteristics than SSGPA. These findings concur with the available literature (e.g. O'Neill et al. 2011; Sinha et al. 2011; Richardson, Abraham, and Bond 2012; Stemler 2012; Cortes 2013; Edwards, Friedman, and Pearce 2013; Kreiter and Axelsson 2013; Gaertner and Larsen McClarty 2015). Note that in The Netherlands, selective university programmes are obliged to use additional admission toolkits besides SSGPA, and that all non-selective university programmes are obliged to offer applicants a study check (e.g. <http://www.uu.nl/bachelors/en/matching>) before commencement.

Of course, SSGPA is an overall measure of both cognitive and non-cognitive factors, but, given the transition from secondary to tertiary education, additional information on academic potential is indeed warranted. Findings that SSGPA is not the best available predictor for academic achievement *after* the first year (Stegers-Jager et al. 2015) add strength to this argument. Theoretically it is logical that including both programme-specific cognitive results (Cortes 2013) and NCF (Schmitt 2012; Stemler 2012) allows students to show their potential regardless of how well they did in secondary education. After all, being a well-functioning student entails more than high grades, and thus an accurate prediction requires more than academic grades alone. Signal detection theory provided relevant additional information supportive of this notion. Having illustrated the opportunities signal detection theory provides to investigate validity in a way that integrates group and individual level in one mixed methods analysis, we now proceed to discuss the information the pay-off matrix provides in evidence-based criteria.

Implementation of signal detection theory

Signal detection theory allowed us to set the cut-off criteria for SSGPA, MS and NCF each to optimise the use of what appeared to be their greatest strength (in terms of optimising the proportion of hits, correct rejections, false alarms and/or misses), and the effects of this on hypothetical selection outcomes were explored. Application of signal detection theory provides admission committees with information on the proportions of the two types of errors that occur; admitting students who will fail and rejecting students who would have been successful. More specifically, MS and NCF appear required in addition to SSGPA to bring the proportion of misses to an acceptable level while maintaining reasonable percentages of admission and correct decisions. The notion that including admission criteria other than SSGPA increases validity was supported more convincingly when applying signal detection theory than when using more conventional statistics.

Although the accuracy of the signal detection theory approach cannot be compared to previous findings, since we did not find any reference to previous applications of signal detection theory in the literature on admission toolkits (but see Sawyer (2013), who does report success rates, accuracy rates and cut-off scores), we did find that this approach averages practically halfway between chance level admission decisions and perfect predictions of academic achievement. Comparing that to 10% explained variance, as far as they can be compared, our comparison of analyses appears to argue in favour of using signal detection theory. The information on the individual level and the receiver operating curves do provide information relevant in decisions on the content and especially criteria of admission toolkits. Signal detection theory is applicable wherever criteria are applied in decision-making regardless of the specific context and goals, for example, in validating criteria of multiple mini interviews (Patterson et al. 2016). Following up on Shulruf, Hattie, and Tumen (2008), Shulruf et al. (2011) and Steenman, Bakker, and Van Tartwijk (2014), future application of signal detection theory might be informative on comparisons of selection outcomes given different operationalisations of SSGPA and academic achievement. The relatively small proportion of unsuccessful students does mean that all interpretations are preliminary and awaiting replication with a larger sample. Replication of our results *across* cohorts, regardless of sample size within cohorts, will also remain a necessity. It should be noted that we were able to do the present analyses because all applicants could be admitted in this cohort; when one is actually selecting applicants the misses and correct rejections cannot be calculated.

Conclusion

Depending on the goals that admission committees aim for by using their admission toolkit (Sternberg et al. 2012; Cortes 2013; Sawyer 2013), one would want to be able to evaluate selection outcomes for both their sensitivity and specificity, and to compare different admission toolkits on both group and individual levels. A signal detection approach to selection of students replicated the validity of SSGPA, MS and NCF as predictors of academic achievement. As a tool to inform future selection or advice by past (hypothetical) selection outcomes, signal detection theory allows policy-makers to make better informed decisions on constructing and evaluating admission toolkits and specific criteria than is possible using the more common correlation/regression statistics. The available information on the relation between predictors/admission tools and academic achievement is extended with the pay-off matrix and the criterion to evaluate selection outcomes on both group and individual levels. Together they give scientists and administrators a lever to push or pull, in order to fine-tune the output of their admission procedure to better match their goals. This will be most profitable if data across cohorts, programmes and different operationalisations of academic achievement are compared.

Acknowledgements

We thank Jutta de Jong for helping, amongst others, with the dozens of figures and Femke van den Brink for her feedback on an earlier draft.

Disclosure statement

The authors declare that they have no conflict of interest.

Notes on contributors

Linda van Ooijen-van der Linden is a teacher and a PhD student at the Clinical Psychology division at the faculty of Social and Behavioural Sciences of Utrecht University. She has extensive experience as a teacher and mentor of freshmen in the Bachelor Psychology and is a member of the admission committee of Bachelor Psychology.

Maarten J. van der Smagt is an associate professor at the Experimental Psychology division at the faculty of Social and Behavioural Sciences of Utrecht University. He teaches at both undergraduate and graduate (MSc and PhD) levels. He is a member of the admission committee of Bachelor Psychology and coordinator of the obligatory first-year Psychology curriculum. He also coordinates the teaching efforts for the Experimental Psychology division. Currently, he is a teaching fellow (2015–2017) with a project to design specific internationally oriented minors. His research focuses on perceptual and cognitive processes.

Liesbeth Woertman was appointed Professor of Quality and Design of Psychology Education in 2009 and is Director Psychology of both the bachelor and the master programmes at the faculty of Social and Behavioural Sciences of Utrecht University. She is the chairperson of the Faculty Club Utrecht University. She is also still involved in teaching clinical psychology in the undergraduate and graduate Psychology curriculum. Self-concept and body-image are her main research topics.

Susan F. te Pas is Associate Dean of Undergraduate Education and Professor of Cognitive Psychology of Higher Education at the faculty of Social and Behavioural Sciences of Utrecht University. She has taught courses in several bachelor and master programmes. At the Psychology department, she was coordinator of the freshman year and initiated the Psychology Freshmen Colleges. She has also coordinated the PhD programme for the Helmholtz Research School and was responsible for the academic master Psychology. In 2011, she became the first teaching fellow of the Faculty of Social and Behavioural Sciences at Utrecht University. Her research covers a wide range of topics within the domain of cognitive psychology and vision science.

References

- Bingham, W. V. 1917. "Mentality Testing of College Students." *Journal of Applied Psychology* 1 (1): 38–45. doi:[10.1037/h0073261](https://doi.org/10.1037/h0073261).
- Christersson, C., D. Bengmark, H. Bengtsson, C. Lindh, and M. Rohlin. 2014. "A Predictive Model for Alternative Admission to Dental Education." *European Journal of Dental Education* 19: 251–258. doi:[10.1111/eje.12129](https://doi.org/10.1111/eje.12129).
- Cortes, C. M. 2013. "Profile in Action: Linking Admission and Retention." *New Directions for Higher Education* 161 (Mar.): 59–69. doi:[10.1002/he.20046](https://doi.org/10.1002/he.20046).
- De Koning, B., S. Loyens, G. Smeets, R. Rikers, and H. Van der Molen. 2012. "Relaties tussen VWO-Eindexamencijfers voor Kernvakken en Studieprestaties in een Bachelorprogramma Psychologie [Relationship Between Secondary School Core Curriculum Grades and College Achievement in Psychology]." *Tijdschrift Voor Hoger Onderwijs* 30 (3): 150–160.
- De Witte, K., and S. J. Cabus. 2013. "Dropout Prevention Measures in the Netherlands, an Explorative Evaluation." *Educational Review* 65 (2): 155–176. doi:[10.1080/00131911.2011.648172](https://doi.org/10.1080/00131911.2011.648172).
- Edwards, D., T. Friedman, and J. Pearce. 2013. "Same Admissions Tools, Different Outcomes: A Critical Perspective on Predictive Validity in Three Undergraduate Medical Schools." *BMC Medical Education* 13: 173. doi:[10.1186/1472-6920-13-173](https://doi.org/10.1186/1472-6920-13-173).
- Ferguson, C. J. 2009. "An Effect Size Primer: A Guide for Clinicians and Researchers." *Professional Psychology: Research and Practice* 40 (5): 532–538. doi:[10.1037/a0015808](https://doi.org/10.1037/a0015808).
- Gaertner, M. N., and K. Larsen McClarty. 2015. "Performance, Perseverance, and the Full Picture of College Readiness." *Educational Measurement: Issues and Practice* 34 (2): 20–33.
- Green, D. M., and J. A. Swets. 1966. *Signal Detection Theory and Psychophysics*. New York: Wiley.
- Huang, C. 2012. "Discriminant and Incremental Validity of Self-Concept and Academic Self-Efficacy: A Meta-Analysis." *Educational Psychology* 32 (6): 777–805. doi:[10.1080/01443410.2012.732386](https://doi.org/10.1080/01443410.2012.732386).
- Kotzee, B., and C. Martin. 2013. "Who Should Go to University? Justice in University Admissions." *Journal of Philosophy of Education* 47 (4): 623–641. doi:[10.1111/1467-9752.12044](https://doi.org/10.1111/1467-9752.12044).
- Kreiter, C. D., and R. D. Axelson. 2013. "A Perspective on Medical School Admission Research and Practice over the Last 25 Years." *Teaching and Learning in Medicine* 25 (suppl. 1): S50–S56. doi:[10.1080/10401334.2013.842910](https://doi.org/10.1080/10401334.2013.842910).
- Macmillan, N. A., and C. D. Creelman. 2005. *Detection Theory: A User's Guide*. 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates.
- O'Neill, L., J. Hartvigsen, B. Wallstedt, L. Korsholm, and B. Eika. 2011. "Medical School Dropout – Testing at Admission versus Selection by Highest Grades as Predictors." *Medical Education* 45 (11): 1111–1120. doi:[10.1111/j.1365-2923.2011.04057.x](https://doi.org/10.1111/j.1365-2923.2011.04057.x).

- Patterson, F., A. Knight, J. Dowell, S. Nicholson, F. Cousins, and J. Cleland. 2016. "How Effective are Selection Methods in Medical Education? A Systematic Review." *Medical Education* 50 (1): 36–60. doi:10.1111/medu.12817.
- Poole, P., B. Shulruf, J. Rudland, and T. Wilkinson. 2012. "Comparison of UMAT Scores and GPA in Prediction of Performance in Medical School: A National Study." *Medical Education* 46 (2): 163–171. doi:10.1111/j.1365-2923.2011.04078.x.
- Preiss, D. D., J. C. Castillo, E. L. Grigorenko, and J. Manzi. 2013. "Argumentative Writing and Academic Achievement: A Longitudinal Study." *Learning and Individual Differences* 28 (Dec.): 204–211. doi:10.1016/j.lindif.2012.12.013.
- Richardson, M., C. Abraham, and R. Bond. 2012. "Psychological Correlates of University Students' Academic Performance: A Systematic Review and Meta-Analysis." *Psychological Bulletin* 138 (2): 353–387. doi:10.1037/a0026838.
- Robbins, S. B., K. Lauver, H. Le, D. Davis, R. Langley, and A. Carlstrom. 2004. "Do Psychosocial and Study Skill Factors Predict College Outcomes? A Meta-Analysis." *Psychological Bulletin* 130 (2): 261–288. doi:10.1037/0033-2909.130.2.261.
- Sartania, N., J. D. McClure, H. Sweeting, and A. Browitt. 2014. "Predictive Power of UKCAT and Other Pre-Admission Measures for Performance in a Medical School in Glasgow: A Cohort Study." *BMC Medical Education* 14 (1): 116–125. doi:10.1186/1472-6920-14-116.
- Sawyer, R. 2013. "Beyond Correlations: Usefulness of High School GPA and Test Scores in Making College Admissions Decisions." *Applied Measurement in Education* 26 (2): 89–112. doi:10.1080/08957347.2013.765433.
- Schmitt, N. 2012. "Development of Rationale and Measures of Noncognitive College Student Potential." *Educational Psychologist* 47 (1): 18–29. doi:10.1080/00461520.2011.610680.
- Shulruf, B., J. Hattie, and S. Tumen. 2008. "The Predictability of Enrolment and First-Year University Results from Secondary School Performance: The New Zealand National Certificate of Educational Achievement." *Studies in Higher Education* 33 (6): 685–698. doi:10.1080/03075070802457025.
- Shulruf, B., Y. G. Wang, Y. J. Zhao, and H. Baker. 2011. "Rethinking the Admission Criteria to Nursing School." *Nurse Education Today* 31 (8): 727–732. doi:10.1016/j.nedt.2010.11.024.
- Simpson, P. L., H. A. Scicluna, P. D. Jones, A. Cole, A. J. O'Sullivan, P. G. Harris, G. Velan, and H. P. McNeil. 2014. "Predictive Validity of a New Integrated Selection Process for Medical School Admission." *BMC Medical Education* 14 (1): 86–95. doi:10.1186/1472-6920-14-86.
- Sinha, R., F. Oswald, A. Imus, and N. Schmitt. 2011. "Criterion-Focused Approach to Reducing Adverse Impact in College Admissions." *Applied Measurement in Education* 24 (2): 137–161. doi:10.1080/08957347.2011.554605.
- Stanislaw, H., and N. Todorov. 1999. "Calculation of Signal Detection Theory Measures." *Behavior Research Methods, Instruments, & Computers* 31 (1): 137–149. doi:10.3758/BF03207704.
- Steenman, S. C., W. E. Bakker, and J. W. F. Van Tartwijk. 2014. "Studies in Higher Education Predicting Different Grades in Different Ways for Selective Admission: Disentangling the First-Year Grade Point Average." *Studies in Higher Education*. doi:10.1080/03075079.2014.970631.
- Stegers-Jager, K. M., A. P. N. Themmen, J. Cohen-Schotanus, and E. W. Steyerberg. 2015. "Predicting Performance: Relative Importance of Students' Background and Past Performance." *Medical Education* 49 (9): 933–945. doi:10.1111/medu.12779.
- Stemler, S. E. 2012. "What Should University Admissions Tests Predict?" *Educational Psychologist* 47 (1): 5–17. doi:10.1080/00461520.2011.611444.
- Stenberg, R. J., C. R. Bonney, L. Gabora, and M. Merrifield. 2012. "WICS: A Model for College and University Admissions." *Educational Psychologist* 47 (1): 30–41. doi:10.1080/00461520.2011.638882.
- Van der Heijden, P. G. M., D. J. Hessen, and T. Wubbels. 2012. "Studiesucces of -Falen van Eerstejaarsstudenten Voorspellen: Een Nieuwe Aanpak [A New Approach to Predicting Freshmen Academic Success or Failure]." *Tijdschrift Voor Hoger Onderwijs* 30 (4): 233–244.
- Visser, K., H. Van der Maas, M. Engels-Freeke, and H. Vorst. 2012. "Het Effect op Studiesucces van Decentrale Selectie middels Proefstuderan aan de Poort [The Effect on Academic Achievement of Local Selection through a 'Work Sample']." *Tijdschrift Voor Hoger Onderwijs* 30 (3): 161–173.