Admission to master programmes: What are the indicators for successful study performance?

Magnus Andersson[1], Johan Karlander[2], Mattias Sandberg[3] and Gunnar Tibert[4]

[1]Dept. Applied Physics, School of Engineering Sciences, KTH
[2]Dept. Computer Science, School of Electrical Engineering and Computer Science, KTH
[3]Dept. Mathematics, School of Engineering Sciences, KTH
[4]Dept. Engineering Mechanics School of Engineering Sciences, KTH

ABSTRACT

Admission of applicants to higher education in a fair, reliable, transparent, and efficient way is a real challenge, especially if there are more eligible applicants than available places and if there are applicants from many different educational systems. Previous research on best practices for admission to master's programmes identified the key question about an applicant's potential for success in studies, but was not able to provide an answer about how to rate the merits of the applicants. In this study, indicators for study success are analysed by comparing the study performance of 228 students in master's programmes with their merits at the time of admission. The null hypothesis was that the applicant's average grade at the time of admission is the only indictor for study success. After testing for potential bias using almost 20 possible other indicators, the null hypothesis had to be rejected for four indicators (in order of importance): (i) university ranking, (ii) length of bachelor's studies within subject, (iii) English language test and (iv) subject matching between bachelor's and master's education. Evaluation of quality of prior education is tricky and results from this study clearly indicate that students from higher ranked universities possess better knowledge and stronger skills for our master's programmes. Work is ongoing to improve the merit rating model by involving more master's programmes at KTH and analysing performance data from a larger number of students.

KEYWORDS

Master, admission, merit rating model, success indicators

INTRODUCTION

Higher education is divided into three cycles: first cycle (bachelor level), second cycle (master level) and third cycle (PhD level). Political initiatives aiming to increase the coherence in higher education systems has led to a collaborative framework called the

European Higher Education Area (EHEA, 2023). One of the aims of this collaboration is to increase mobility between universities and to facilitate for students to obtain a bachelor's degree from one university and a master's degree from another university in Europe. With such a system, it is also possible for non-European students to directly apply for a master's programme in Europe.

However, with applicants coming from different educational systems with different curricula and different educational traditions, it is a substantial challenge to decide about admission or non-admission in a fair, reliable, and transparent way. This is even more challenging if there are more eligible applicants than available places on a master's programme. In such a case, it is necessary to rank the eligible applicants using a merit rating model. Important work addressing best practices for master's level admission was performed in the EU funded MasterMind Europe project (MasterMind Europe, 2023). This effort set up guiding tools within six different areas, which were labelled (Mastermind Europe, 2023):

1) Coherent admission framework
2) Subject-related knowledge & skills
3) General academic competence
4) Personal competencies & traits
5) Language requirements
6) Managing graduate admission

The strength of the MasterMind Europe project was that it set up a framework for the administration of the admission process. However, it was not able to create consensus on best practices for merit rating models and the guiding tool "1) Coherent admission framework" (MasterMind Europe, 2023) only identified relevant questions, but was not able to give any answers. However, an important step forward in this work was to identify that the key question for admission is "*Does this applicant have 'what it takes' to be successful in our master's programme?"* (MasterMind Europe, 2023, Guiding tool 1: Coherent admission framework, p.16). Hence, a large variability in admission practices and models between different master's programmes still remains (Chari & Potvin 2019; MasterMind Europe, 2023). It has been argued that this is partly due to shortcomings in the common standardization and recognition approach still used in the admission process (Kouwenaar, 2015).

An alternative and much less investigated approach to validate merit rating models is to look at study performance of students within a master's programme and correlate this with data available at the time of admission. In one of few studies, Zimmermann et al. (2015) analyse study performance of students coming from a bachelor's programme in Computer Science and entering a master's programme in Computer Science at the same university. They analysed 81 variables (size of data set, N = 171) using linear regression and found that the strongest indicator of study success at master's level (measured as GPA, grade point average) was GPA obtained during the last (third) year of the bachelor's programme and that overall GPA was still an important factor (Zimmermann et al., 2015). In another study of the admission to a master's programme in Data Science, Zhao et al. (2020b) tested different classification algorithms for predicting study performance (size of data set, N = 132). They reported a good predictive power to identify high and low performing students and found that the GPA from the bachelor's degree has a strong impact on performance as does the previous major (Zhao et al, 2020b). A machine learning approach has also been tested on this problem (Zhao et al, 2020a), but such an approach has the disadvantage of not being sufficiently transparent.

In this work, we will investigate how admission data can be used to predict study performance at a master's programme and, hence, lead to better validity of the merit rating model. The focus is to identify valid factors and to get a rough estimate about how important they are for student success.

METHODOLOGY

In this work, we have used performance data from N = 228 students admitted during the years 2018-2020 and later enrolled in one out of seven master's programmes at the School of Engineering Sciences at KTH. Performance data was found in Ladok, which is a Swedish national system to assists higher education institutions in Sweden to document study performance. Letter grades (*A*-*E* with *A* as the highest grade) given for passed courses at our university were transformed into numerical values according to the following scheme:

$$A = 5.0 \;\; ; \;\; B = 4.5 \;\; ; \; C = 4.0 \; ; \; D = 3.5 \; ; \; E = 3.0$$

An average grade was then calculated for all passed courses during the nominal study time of two years for a master's programme and the total number of ECTS credits passed after two years was also calculated. To facilitate data analysis and only have one performance parameter to consider in our analysis, we define a single overall performance value (PV) according to the following equation

$$PV = \frac{Average\ grade - 1}{4} \bullet \frac{Credits\ passed}{Expected\ credits}$$

As a starting point, we compare the performance value with a merit rating that is only based on average grade in the previous bachelor's degree. We then use statistical analysis to determine what other indicators create bias in the data and, therefore, are also valid indicators of study performance. Finally, by comparing study performance during master's studies at KTH to different variants of a linear model for merit rating, we can optimize the merit rating model and estimate the relative importance of different indicators for predicting study performance.

RESULTS

Fundamental data for all students in our analysis are shown in Fig. 1. Each point corresponds to one student and compares the average grade from bachelor's level (on the abscissa) with performance value at master's level (on the ordinate). As described above, the performance value is a weighted value based both on average grade at master's level and expected number of credits that a student is expected to have passed during two years of studies. As an example, a student that have passed all courses within two years with the lowest possible pass grade (*E*) in all courses, has a performance value of 0.5. Students with a performance value below 0.5 has not finished all their courses after two years. The red line corresponds to a direct relation between average grades at bachelor's level and performance value at master's level. Students lying above (below)

the line perform better (worse) at master's level than predicted from their average grade at bachelor's level.

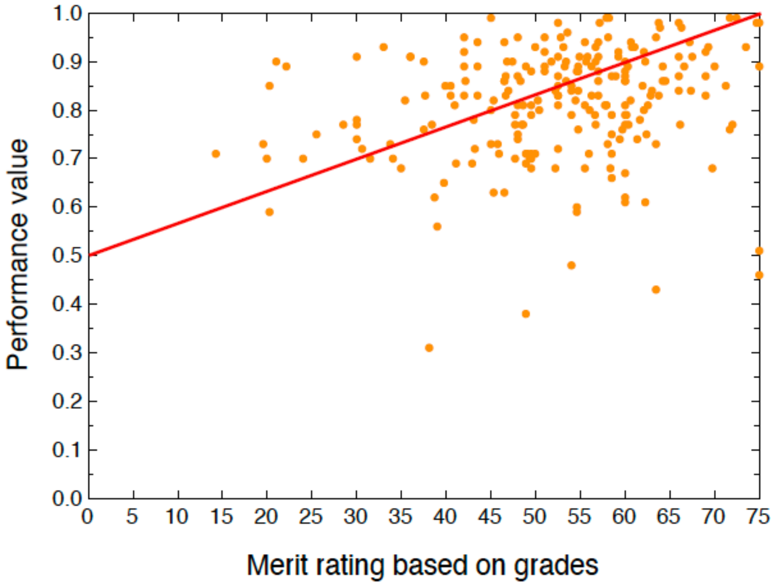

Figure 1. Student performance value at master's level (average grades and passed credits) versus average grade at bachelor's level. The scale on the abscissa is from 0 (lowest pass grade on all bachelor courses) to 75 (highest pass grade on all bachelor courses).

With the graph shown in Fig. 1 as a starting point, it is possible to systematically look for the importance of other indicators for study success at master's level. For each possible indicator available at admission, student data is divided into two groups to look for bias in the data. A Wilcoxon rank-sum test (Wilcoxon, 1945; Mann & Whitney, 1947) is performed to statistically determine the probability that the two groups have the same distribution. Hence, we apply a null hypothesis that data points in the two groups will deviate from the line in Fig. 1 with the same probability distribution of lying above or below the line. An example of such an analysis is shown in Fig. 2, where we have divided the data points from Fig.1 into two categories based on university ranking. Students with a bachelor's degree from universities at ranking positions higher than 400 on both the QS (QS Top universities, 2023) and the THE (THE Times Higher Education, 2023) university ranking lists are marked with blue points. It is directly seen in Fig. 2 that there is a clear bias in data and that students from those universities perform worse at KTH than students from universities ranked on positions between 1-400 on either of these ranking lists.

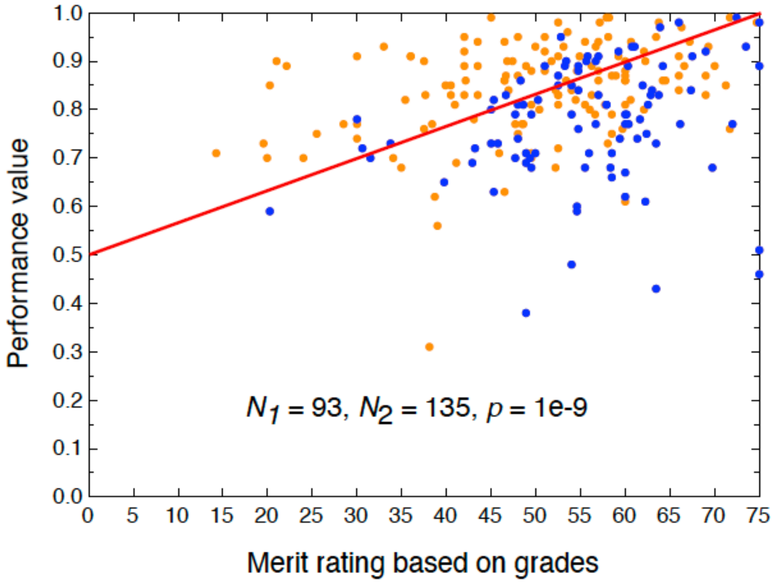

Figure 2. Bias related to university ranking. Students with bachelor's degree from universities at ranking positions higher than 400 on both the QS and THE university ranking lists are shown as blue points (N_1 = 93). Students from higher ranked universities (N_2 = 135) are shown in orange.

The size of our data set allowed us to test almost 20 possible indicators and the null hypothesis of equal distribution had to be rejected for the following indicators (an upper limit of the probability of equal distribution is given within parenthesis):

• University ranking > 400 ($p < 10^{-8}$)

• 4-year bachelor's degree from a Spanish university ($p < 10^{-4}$)

• Just passed general eligibility in English ($p < 10^{-3}$)

• Good matching between name of bachelor's and master's education ($p < 10^{-2}$)

Hence, these are all valid indicators and need to be considered in some way in a merit rating model. Also, our data was not able to statistically show a difference for factors such as English knowledge above the minimum eligibility level, gender or choice of priority when applying to several programmes. A limitation to our study is that we could only test indicators used earlier in the merit rating models at the School of Engineering Sciences at KTH, which means that our analysis may have missed other valid indicators for successful studies at a master's programme.

Once the valid indicators for study success have been identified, the next step is to find out how to weigh these indicators in the merit rating model. A more precise answer to that question requires more data, but we tried to get a rough idea about this from our limited amount of data. We assumed that a merit value can be described as a weighted linear combination of grades (*Grade*), university (*Uni*) and English proficiency (*Eng*). The reason for not including the other two valid indicators (4-year bachelor's degree from Spanish university and good matching between name of bachelor and master education) was partly based on the difficulty to parametrize these indicators and partly due to the limited amount of data available. Data for the three indicators in the model were first scaled to roughly lie in the interval between 0 and 25, with 0 denoting the lowest possible score (lowest passing grade in every bachelor course, lowest university ranking, and just passing the general eligibility in English), and 25 denoting the maximum possible score in the three domains. We then formed a merit rating value M given by the following weighted average

$$M = k_{!} \bullet Grade + k_{"} \bullet Uni + k_{\#} \bullet Eng$$

where the coefficients $k_{!}$, $k_{"}$ and $k_{\#}$ must all be positive. Furthermore, we required that the sum of the coefficients $k_{!} + k_{"} + k_{\#} = 3$, to ensure that we obtain a total merit rating value in the interval between 0 and 75 (the standard range used for our merit rating). We then look for an optimal set of coefficients by discretizing the set they belong to in the following way. Each of the coefficients are allowed to only take the discrete values 0, 0.01, 0.02, …, 3, and their sum must still be three. For each of the possible choices of parameters we perform a linear regression of the study performance value versus the composite merit rating value M. The coefficients that give the largest slope of the regression linear function, and whose corresponding merit rating value hence has the greatest effect on study performance, was found to be $k_{!} = 1.2$, $k_{"} = 1.2$ and $k_{\#} = 0.6$. Hence, grades and university parameters should be given about equal weight, while the test of English proficiency should be given about half the weight of those two parameters. Fig. 3 shows the linear regression using the optimal parameters.

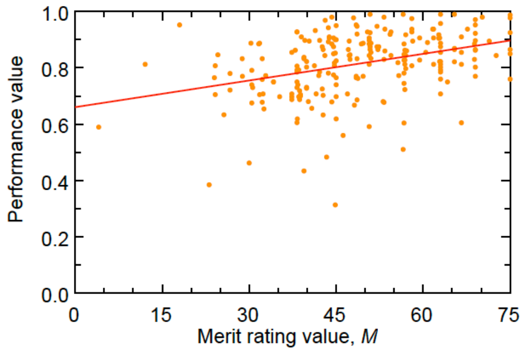

Figure 3. Linear regression fit of performance value versus the merit rating value computed with the optimal coefficients $k_{\#} = 1.2$, $k_{\$} = 1 - 2$ and $k_{\%} = 0.6$ using a linear model. The linear function in red has the equation: $PV = 0.0032 \bullet M + 0.66$.

DISCUSSION

From an evaluation point of view, fair evaluations of applicants require both validity (the use of relevant parameters) and reliability (the correct and accurate measures of these parameters). In addition, the evaluation process must be efficient in order to keep down the use of resources, but that aspect lies outside of the scope of this work. Our statistical analysis has revealed four parameters as valid indicators for student success at a master's programme. We will now discuss the fundamental reasons for the connection between these indicators and student success and their reliability in a merit rating model.

University ranking was found to be about equally important as average grade in predicting student success at our master's programmes. There are several reasons for this. Firstly, students compete with their grades when entering university and almost everywhere, high average grades are required to be admitted to a highly ranked university. This implies that there is already at entrance a significant difference in background knowledge and skills of the cohorts of students admitted to different universities. Faculty at a highly ranked university then start the education with a student cohort having higher average grades from the beginning. If faculty members do their work properly, this difference will persist and students graduating with a bachelor's degree from a highly ranked university are expected to have acquired more demanding knowledge and skills than students graduating from lower ranked universities. Secondly, there are also differences in university culture and faculty engagement in research frontiers. University ranking depends to a large extent on the faculty's research output and university reputation, which to a large extent is what students are looking at when applying to an education. Hence, university ranking based on reputation and faculty engagement in research approximates the background knowledge and skills of students when entering university. However, it does not consider that universities may offer different learning efficiency for their students. Hence, due to its indirect measure of student ability, there are clearly concerns about the reliability of using university ranking for the prediction of student success at a master's programme. In fact, one would like to develop more reliable ways than university ranking to measure the indisputable differences between educations at different universities.

The second most important indicator that falsified the null hypothesis was one additional year of studies within the subject (which is the case for Spanish students with a 4-year bachelor's education). This is not at all surprising, since one addition year of study within a subject will give a more profound knowledge base and students entering a master's education with such a background have already acquired some of the knowledge taught in the master's programme. This makes it considerably easier for them to adjust to the master's programme. However, it is difficult to parametrize this indicator in a reliable way.

The third indicator for successful studies at a master's programme was knowledge of English that lies above the minimum requirements. Having just passed the minimum level of English proficiency probably makes it harder for a student to assimilate the education. In fact, a marginal understanding of English creates an additional cognitive load on the student. When this additional cognitive load adds to the normal cognitive load in learning a subject, it is not surprising that that the additional cognitive load leads to lower study performance. The indicator for English proficiency can be reasonably well parametrized through standardized English tests, which gives a relatively good reliability. The challenge is to compare between different English tests and to parametrize English proficiency proved by grades in previous English courses.

The arguments put forward above about better study performances for students that have spent more time within the subject also holds for the last valid indicator, which is a comparison between the main curriculum at bachelor's level and master's level. It is difficult to parametrize this indicator in a reliable way and more research is needed. However, neither of the indicators that our data show to be valid for the merit rating model is a complete surprise.

All the indicators listed above are valid indicators of student performance at a master's programme and should be considered when building a merit rating model. However, there may be other valid indicators and there are still many unresolved questions related to reliability and how to parametrize the indicators.

FUTURE WORK

A larger amount of performance data will obviously give more precise answers to many of the questions raised in this work about how to find an optimal, valid and reliable merit rating model that can predict student success at a master's programme. In fact, we are currently working on a project that involves more than twice as many master's programmes and more than ten times as many performance data. We expect to soon be able to present the results of this extended study.

ACKNOWLEDGEMENTS

We would like to thank administrative master coordinator Molea Jung for contributing to the discussions.

REFTERENCES

Chari, D., & Potvin. G. (2019). Admission practices in terminal master's degree-granting physics departments: a comparative analysis. *Physical Review Physics Education Research* 15, 010104.

EHEA (2023, August 10). *European Higher Education Area and Bologna Process*. EHEA. https://www.ehea.info

Kouwenaar, K. (2015). Changing paradigms: towards competency-assessment in admission to master's programmes in Europe – a review. *Turning Journal of Higher Education* 3(1), 99-135.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics* 18(1), 50-60.

Mastermind Europe (2023, April 8). *Tools and Reports*. Mastermind Europe. http://mastermindeurope.eu/tools-and-reference-material/#downloads

QS Top Universities (2023, January 10). *QS World University Rankings 2023: Top global universities*. QS Quacquarelli Symonds Limited. https://www.topuniversities.com/university- rankings/world-university-rankings/2023

THE Times Higher Education (2023, January 10). *World University Rankings 2023*. THE Times Higher Education. https://www.timeshighereducation.com/world-university-rankings/2023/world- ranking

Wilcoxon, F. (1945). Individual comparison by ranking methods. *Biometrics Bulletin* 1, 80-83.

Zhao, Y., Lackaye, B., Dy, J. G., & Bodley, C. E. (2020a. A quantitative machine learning approach to master student admission for professional institutions. In: *Proceedings of the 13th International Conference on Educational Data Mining (EDM 2020)*, A. N. Rafferty, J. Whitehall, V. Cavalli-Sforza & C. Romera (eds.), 538-544.

Zhao, Y., Xu, Q., Chen, M., & Weiss, G. M.. (2020b). Predicting student performance in a master of data science program using admission data. In: *Proceedings of the 13th International Conference on Educational Data Mining (EDM 2020)*, A. N. Rafferty, J. Whitehall, V. Cavalli- Sforza & C. Romera (eds.), 325-333.

Zimmermann, J., Brodersen, K. H., Heinimann, H., & Buhmann, J. M. (2015). A model-based approach to predicting graduate-level performance using indicators of undergraduate-level performance. *Journal of Educational Data Mining* 7(3), 151-176.

ABOUT THE AUTHORS

This section starts on a new page. Introduce each author in a paragraph. Provide full contact details for the corresponding author. To be included in the proceedings, the authors must grant the Creative Commons version 4.0 license (see the link below for explanation).

Magnus Andersson has a PhD in Materials Physics and holds a position as associate professor at the *Department of Applied Physics* at KTH. Since 2017, he is the director of the master's programme in Engineering Physics at KTH. He has worked part time as pedagogical developer and has during the last years published several research papers related to educational research.

Johan Karlander has a PhD in Mathematics and holds a position as associate professor at the *Department of Theoretical Computer Science* at KTH. He has previously worked as programme director. His main research areas are Combinatorics and Game Theory.

Mattias Sandberg has a PhD in Mathematics and holds a position as associate professor at the *Department of Mathematics* at KTH. Since 2021, he is the director of the master's programme Applied and Computational Mathematics at KTH. His main research areas are numerical methods for optimal control problems, ab initio molecular dynamics, and neural networks.

Gunnar Tibert has a PhD in Structural Mechanics and holds a position as associate professor at the *Department of Engineering Mechanics* at KTH. Since 2013, he has been appointed director of first and second cycle studies at the School of Engineering Sciences at KTH and since 2020, he is deputy head of the School of Engineering Sciences. His main research activities have been on design and testing of structures for space applications.

CORRESPONDING AUTHOR

Magnus Andersson KTH

Department of Applied Physics 106
91 Stockholm magnusan@kth.se