CrossMark

A multi-site study on medical school selection, performance, motivation and engagement

A. Wouters[1,2] · G. Croiset[1,2] · N. R. Schripsema[3] ·
J. Cohen-Schotanus[3] · G. W. G. Spaai[4] · R. L. Hulsman[4] ·
R. A. Kusurkar[1,2]

Abstract Medical schools seek ways to improve their admissions strategies, since the available methods prove to be suboptimal for selecting the best and most motivated students. In this multi-site cross-sectional questionnaire study, we examined the value of (different) selection procedures compared to a weighted lottery procedure, which includes direct admission based on top pre-university grade point averages (≥ 8 out of 10; top-pu-GPA). We also considered whether students had participated in selection, prior to being admitted through weighted lottery. Year-1 (pre-clinical) and Year-4 (clinical) students completed standard validated questionnaires measuring quality of motivation (Academic Self-regulation Questionnaire), strength of motivation (Strength of Motivation for Medical School-Revised) and engagement (Utrecht Work Engagement Scale-Student). Performance data comprised GPA and course credits in Year-1 and clerkship performance in Year-4. Regression analyses were performed. The response rate was 35% (387 Year-1 and 273 Year-4 students). Top-pu-GPA students outperformed selected students. Selected Year-1 students reported higher strength of motivation than top-pu-GPA students. Selected students did not outperform or show better quality of motivation and engagement than lottery-admitted students. Participation in selection was associated with higher engagement and better clerkship performance in Year-4. GPA, course credits and strength of motivation in Year-1 differed between students admitted through different selection procedures. Top-pu-

---

✉ A. Wouters
a.wouters@vumc.nl

1 Research in Education, VUmc School of Medical Sciences, VU University Medical Center, PK KTC 5.002, Post box 7057, 1081 BT Amsterdam, The Netherlands

2 LEARN! Research Institute for Learning and Education, Faculty of Psychology and Education, VU University Amsterdam, Amsterdam, The Netherlands

3 Center for Research and Innovation in Medical Education, UMC Groningen, Groningen, The Netherlands

4 Center for Evidence-Based Education, AMC-UvA, Amsterdam, The Netherlands

GPA students perform best in the medical study. The few and small differences found raise questions about the added value of an extensive selection procedure compared to a weighted lottery procedure. Findings have to be interpreted with caution because of a low response rate and small group sizes.

Introduction

By applying selection procedures, medical schools aim to admit motivated students who will perform well in their studies (Turner and Nicholson 2011). However, the currently available selection tools, which are usually combined in selection procedures, appear to be suboptimal for identifying the most suitable candidates. While academic records, multiple mini-interviews (MMIs), aptitude tests, situational judgement tests and selection centres are among the most promising selection tools, none of these have proven to be perfect in terms of reliability, validity, fairness and cost-effectiveness (Cleland et al. 2012; Patterson et al. 2016). Evidence for the added value of costly selection procedures compared to a weighted lottery procedure, which was applied in the Netherlands for many years, is not unequivocal. The literature on selection mainly contains single-site studies and studies investigating selection tools in isolation, rather than combinations of selection tools (Patterson et al. 2016). This multi-site study aims to fill these gaps by examining the value of (different) selection procedures compared to a weighted lottery procedure, which is weighted for pre-university grade point average (pu-GPA) and includes direct admission for students with top-pu-GPAs (≥ 8 out of 10) (Ten Cate 2007). Outcomes of interest were student performances, as well as motivation (for studying medicine) and engagement in learning, because these variables are deemed important for the learning, performance and well-being of students (Casuso-Holgado et al. 2013; Prins et al. 2009; Williams et al. 1999). Motivation concerns the reasons people act in certain ways. These reasons can originate from within the person or from external factors. According to the Self-determination theory (SDT), autonomous motivation (AM) is seen when one does something out of genuine interest or because of a positive valuation of the activity. AM is associated with better learning, academic performance and well-being, compared to controlled motivation (CM), which is seen when one experiences internal or external pressure (Artino et al. 2010; Kusurkar et al. 2011a, 2013; Moulaert et al. 2004; Ryan and Deci 2000; Sobral 2004; Stegers-Jager et al. 2012; Vansteenkiste et al. 2005; Williams et al. 1999). Student engagement also contributes to better learning and academic performance of (medical) students (Carini et al. 2006; Casuso-Holgado et al. 2013; Schaufeli et al. 2002a; Svanum and Bigatti 2009) and has a negative relationship with burnout (Schaufeli et al. 2002b). Engagement is defined as "*a positive, fulfilling, and work-related state of mind that is characterized by vigour, dedication, and absorption*" (Schaufeli et al. 2002b).

The effects of the various admission pathways (i.e., admission based on a selection procedure, a weighted lottery procedure and top-pu-GPA) have been studied and the results are inconclusive. Whenever performance differences are found, however small they are, applying a selection procedure seems favourable over applying weighted lottery (de Visser et al. 2016; Lucieer et al. 2015; Schripsema et al. 2014; Urlings-Strop et al.

2009, 2011, 2013). However, the differences are often not statistically significant (de Visser et al. 2016; Hulsman et al. 2007; Lucieer et al. 2015; Schripsema et al. 2014; Stegers-Jager et al. 2015; Urlings-Strop et al. 2009). A more consistent finding is that top-pu-GPA students generally outperform selected and lottery-admitted students (de Visser et al. 2016; Schripsema et al. 2014). Different contexts of the single-site studies (such as selection procedures, proportion of students admitted through the different admission pathways) may have led to conflicting results. Moreover, a variety of outcome measures, mainly pertaining to the pre-clinical phase, has been used. Research on the motivation, especially quality of motivation, of students admitted through different pathways is scarce and the findings show a similar pattern. Either no significant differences (Nieuwhof et al. 2004; Wouters et al. 2016), or better quantity and quality of motivation among selected students have been reported (Hulsman et al. 2007; Kusurkar et al. 2010, 2013; Wouters et al. 2016). Despite its implications for learning and performance, engagement has not yet been studied as an outcome measure of selection. The first of three research questions addressed in this study is: Are different admission groups associated with differences in motivation, engagement and pre-clinical and clinical performance? Based on the literature, we hypothesize that top-pu-GPA students will outperform selected and lottery-admitted students, and selected students will outperform lottery-admitted students in pre-clinical and clinical education, while selected students will outperform and report higher AM and engagement than lottery-admitted and top-pu-GPA students.

Some researchers have made use of the fact that the lottery-admitted group consists of two types of students: students who had participated in selection and students who refrained from it. It has been argued that applicants who invest the time and effort necessary for participation in selection may perform better than those who refrain from it (Schripsema et al. 2014), suggesting that selection may attract a group of better quality students. The evidence for this is scarce and findings are inconclusive. Students who had not participated in selection have been found to underperform compared to students who were selected or students who had enrolled through weighted lottery after being rejected in selection (de Visser et al. 2016; Schripsema et al. 2014), but these findings did not always reach significance (de Visser et al. 2016; Schripsema et al. 2014; Urlings-Strop et al. 2013). Thus, limited evidence supports the hypothesis that students who have participated in selection outperform those who have not. Motivation, as reflected in the preparation for selection, has been suggested as one reason why students who have participated might perform better (Schripsema et al. 2014, 2016; Wouters et al. 2016), but this assumption has not yet been investigated. The second research question addressed in this study is: Is participation in selection associated with differences in motivation, engagement and pre-clinical and clinical performance? Based on the literature, we hypothesize that students who participated in selection outperform and report higher AM and engagement than students who did not participate in selection.

Because selection procedures are usually costly, identifying which type of procedure is associated with the most desirable student characteristics can inform policy decisions. However, the mere presence of an effort-intensive selection procedure may be more important than the specific characteristics of the procedure. The single-institution nature of previous research has resulted in a relative lack of studies comparing different selection procedures. A study comparing students selected using cognitive criteria and students selected using non-cognitive criteria within one medical school, showed no differences with regard to dropout and performance during the pre-clinical and clinical phases of medical study (Lucieer et al. 2015). The present study includes multiple institutions applying different selection procedures, enabling comparisons across medical schools. The third research question addressed in this study is: Are

different selection procedures associated with differences in motivation, engagement and pre-clinical and clinical performance? Based on the limited evidence and the notion that the content of the selection procedure may be secondary to the presence of a selection procedure, we hypothesize that different types of selection procedures do not result in differences in motivation, engagement and pre-clinical and clinical performance.

Methods

Study design

This was a multi-site cross-sectional study using an online survey (Net Questionnaire) comprised of personal data and standard, validated questionnaires. The indicators of academic performance of the participating students were retrieved from student administrative databases.

Setting

This study was carried out at three of the eight Dutch medical schools: VUmc School of Medical Sciences Amsterdam (VUmc), Academic Medical Center Amsterdam (AMC), and University Medical Center Groningen (UMCG). Inclusion of these medical schools was based on differences in selection procedures and their use of selection at least since 2010 (for enabling inclusion of Year-4 performance). Medical study in the Netherlands consists of three years of pre-clinical education, followed by three years of clinical education, after which students obtain their medical degrees. Although small local differences may exist, we expect the curricula to be largely comparable because medical curricula across the Netherlands are all vertically integrated, student-centred (Ten Cate 2007) and driven by nationally standardized end terms (Van Herwaarden et al. 2009). An overview of the characteristics of the selection procedures of the different medical schools is provided in Table 1.

Participants

In the 2013–2014 academic year, students were invited via e-mail (with two reminders) to participate in this study. Participation was voluntary and in the e-mail, students were informed about the aims of the study and handling of data. At the beginning of the survey, students gave their informed consent. The sample consisted of students from Year-1 (pre-clinical phase) and Year-4 (clinical phase) because assessment in Year-1 is based on cognitive skills and assessment in Year-4, the first clinical year of the study, is mostly based on non-cognitive skills. Moreover, the selection procedures may be associated with preclinical and clinical performance to a different extent. For every ten participants, a gift card of €25 was awarded through random selection.

Outcome measures

Academic performance

Three measures were defined to represent academic performance in Year-1: course credits, GPA and professional behaviour. *Course credits* (European credits) obtained in the respective study year were used. At all medical schools, the maximum number of course

Table 1 Differences of the selection procedures of the three universities

	VUmc	AMC	UMCG
Selection procedure	*Procedure A* Two phases: 1. Portfolio including previous academic records and extracurricular activities. Students meeting the set threshold were invited to participate in the second phase	*Procedure B* Two phases: 1. Cognitive tests and portfolio including previous academic records and extracurricular activities	*Procedure C* Two phases: 1. Portfolio (comprising sections on pre-university education, extracurricular activities, and reflection) and academic and non-academic tests The highest scoring applicants were invited to participate in the second phase
	2. Lectures followed by assessment of academic skills, measured with tests about medical subjects and study skills	2. Lecture followed by an academic test and three-station MMI (Year-1) or interview (Year-4)	2. Patient lecture followed by assignments (related to the lecture, writing an essay, and scientific reasoning) and a four-station MMI assessing communication skills, collaboration skills and reflection
	(http://www.med.vu.nl/nl/opleidingen/bachelor-geneeskunde/decentrale-selectie/index.aspx)	(https://www.amc.nl/web/Onderwijs/Aankomend-student/Geneeskunde/Decentrale-selectie-1.htm)	(http://www.rug.nl/umcg/education/medicine/selection_-admission-requirements-and-deficiencies)
Places assigned through selection	Year-1: 60% of 350 places Year-4: 50% of 350 places	Year-1: 75% of 350 places Year-4: 50% of 350 places	Year-1: 100% of 410 places Year-4: 50% of 410 places

credits per year was 60. *GPA* in the first year was comprised of the average of the first attempts on all knowledge tests. For VUmc, AMC and UMCG, respectively, six, five, and four tests were included. For *professional behaviour,* unsatisfactory, satisfactory or good judgments on professional development were examined. Similar to other researchers, we chose the achievement of *good clerkship performance* as an indicator of performance in Year-4 (Stegers-Jager et al. 2015). Good clerkship performance was defined as receiving a grade of 8 or higher out of 10 for at least half of the clerkships. Clinical educators tend to be reluctant to fail students for their clerkships (Daelmans et al. 2016) and clerkship grades are usually above average. By using good clerkship performance as an outcome measure, we aimed to identify the students that stood out by performing well in the majority of their clerkships. The final clerkship grade is a single grade that includes an assessment of professional behaviour. Year-4 was composed of six clerkships at VUmc and AMC and four clerkships at UMCG.

Motivation

Two measures were defined to represent motivation: strength of motivation and type of motivation (autonomous and controlled; AM and CM). We used the concept of motivation put forth by Self-determination Theory (SDT) (Deci and Ryan 1985). AM and CM were measured with the 16-item Academic Self-regulation Questionnaire (Vansteenkiste et al.

2009). Scores ranged from 1 (not important at all) to 5 (very important). Example items are "I am studying medicine because I want to learn new things" and "I am studying medicine because I want others to think I'm smart." for autonomous and controlled motivation, respectively. Relative autonomous motivation (RAM) was calculated by subtracting the CM subscale score from the AM subscale score. Strength of motivation was measured with the 15-item Strength of Motivation for Medical School-Revised questionnaire (SMMS-R; Kusurkar et al. 2011b; Leibach and Stern 2013; Nieuwhof et al. 2004). An example item is "I would still choose medicine even if that meant I would never be able to go on holidays with my friends anymore". Scores ranged from 1 (strongly disagree) to 5 (strongly agree).

Engagement

The total score on the nine-item Utrecht Work Engagement Scale-Students (UWES-S-9) (Schaufeli et al. 2002a) represented students' engagement. Students rated their level of engagement across the domains of vigour, absorption and dedication. An example item is "When I am studying, I forget everything else around me". Scores ranged from 0 (never) to 6 (always).

Independent variables

To answer the three research questions, three independent variables were defined: *admission group* (selection, lottery, and top-pu-GPA)*, selection participation* (participation and no participation in selection) and *selection procedure* (selection procedures A, B and C for the selection procedures at VUmc, AMC and UMCG, respectively).

Confounders

We investigated whether the variables age, gender, university, first-generation student, doctor parent, ethnicity, area of growing up, living situation and pu-GPA needed to be included as confounders in the final models, along with the independent variables. Ethnicity was defined using the definition of Statistics Netherlands (CBS; www.cbs.nl), which states that a person belongs to an ethnic minority group if at least one of his or her parents was born outside the Netherlands. These variables were indicated as possible confounders because previous research showed the importance of students' background characteristics in performance and motivation (Kusurkar et al. 2011a; Stegers-Jager et al. 2015; Strauser et al. 2012).

Statistical analysis

For linear and dichotomous outcome variables, respectively, linear and binary logistic regression modelling was performed. First, we performed univariate regression analyses. Next, for every regression model, we investigated whether variables needed to be included as confounders in the final model based on a change in the regression coefficient of 10% or more and a significant association with the outcome variable (Twisk 2006). Wherever appropriate, we used the Bonferroni procedure for multiple comparison correction. Pu-GPA was not considered as a possible confounder in the analyses in which the different admission groups (top-pu-GPA, selection and lottery) were compared, because the top-pu-GPA group was, by definition, the group with the highest pu-GPAs. Analyses were

performed using IBM SPSS Statistics for Windows Version 20.0 (IBM Corp., Armonk, NY, USA).

Results

First, we provide the descriptives and the reliability tests of the used scales. Next, we report the results for each research question separately.

The 666 participants (response rate $\approx 35\%$ across all three universities) included 387 Year-1 students and 273 Year-4 students. The average ages of the participants (18.7, 19.1 and 18.5 for Year-1 students and 23.2, 22.9 and 22.6 for Year-4 students from VUmc, AMC, and UMCG students) fairly reflected the average ages of their respective cohorts (19.3, 19.2 and 18.4 for Year-1 students and 24.2, 23.5 and 23.0 for Year-4 students from VUmc, AMC and UMCG). Female students were slightly overrepresented in our study sample (77.5, 68 and 78% for Year-1 students and 78.8, 73.3 and 72.4% for Year-4 students from VUmc, AMC and UMCG) compared to the respective cohorts (67.3, 59.2 and 78% for Year-1 students and 68.1, 68.7 and 67.2% for Year-4 students from VUmc, AMC and UMCG). Participants who enrolled in a graduate entry programme (n = 6) and participants admitted under special circumstances (n = 17) were excluded from the analyses. Seventy-six students (12%) were admitted based on top-pu-GPA, 75 students (12%) enrolled through a weighted lottery without having participated in a selection procedure, 82 students (13%) were admitted through a weighted lottery after being rejected in selection and 395 students (61%) were admitted through selection. We considered the admission pathway distribution in our sample to be a fair reflection of the population, based on the places assigned through selection at each medical school. Of the Year-1 participants 62.5, 70.3 and 92.4% at VUmc, AMC and UMCG, respectively, were admitted through selection. Of the Year-4 participants 44.2, 45.0 and 40% at VUmc, AMC and UMCG, respectively, were admitted through selection. A further breakdown by study year is provided in Supplementary File 1 (Appendix).

The Crohnbach's alpha values for reliability for the UWES-S-9, AM, CM and the SMMS-R were 0.90, 0.82, 0.84 and 0.79 respectively.

Table 2 summarizes the means and standard deviations of all variables. The results of the regression analyses for Year-1 and Year-4 students and Pearson correlations between linear and dichotomous variables are depicted in Tables 3 and 4 and Supplementary File 1 (Appendix), respectively. The incidence of unsatisfactory judgments for professional behaviour was too low (1.4%) to conduct further analyses.

Admission group

Students with top-pu-GPAs obtained higher GPAs (B = 0.526, $p < 0.01$) and were more likely to show good performance during their clerkships [Odds Ratio (OR) 1.218, $p < 0.1$] than selected students. Selected students reported higher strength of motivation in Year-1 than students with top-pu-GPAs (B = 2.581, $p < 0.05$, respectively). Analyses showed no significant associations between admission group and course credits, engagement, strength of motivation in Year-4, AM, CM and RAM. These findings partly support our hypothesis that top-pu-GPA students would outperform other students, while selected students would report higher AM and engagement than top-pu-GPA and lottery-admitted students. Top-pu-GPA performed best, and selected students reported higher strength of motivation, but

Table 2 Descriptives

	Year-1										Year-4								
	n	Female (%)	Age Mean (SD)	ENG Mean (SD)	AM Mean (SD)	CM Mean (SD)	RAM Mean (SD)	SMMS Mean (SD)	GPA Mean (SD)	Course credits Mean (SD)	n	Female (%)	Age Mean (SD)	ENG Mean (SD)	AM Mean (SD)	CM Mean (SD)	RAM Mean (SD)	SMMS Mean (SD)	Good clerkship performance (%)
Participation																			
No	16	81.3	21.0 (3.4)	4.4 (1.0)	4.2 (0.5)	2.0 (0.4)	2.3 (0.7)	53.4 (7.5)	6.5 (0.9)	51.0 (15.3)	59	67.8	23.5 (2.6)	3.7 (0.9)	4.1 (0.5)	2.0 (0.8)	2.1 (1.1)	51.6 (7.6)	67.9
Yes	315	76.8	18.8 (1.3)	4.3 (0.8)	4.3 (0.4)	2.0 (0.7)	2.3 (0.8)	55.5 (6.9)	7.1 (1.1)	56.5 (9.4)	165	77.0	22.6 (1.6)	4.1 (0.9)	4.2 (0.4)	1.9 (0.7)	2.3 (0.8)	52.3 (6.2)	85.9
Selection procedure																			
A	75	76.0	18.2 (0.6)	4.4 (0.8)	4.3 (0.5)	1.9 (0.6)	2.4 (0.8)	54.2 (7.3)	6.5 (0.9)	54.9 (11.3)	23	82.6	23.1 (2.4)	4.3 (0.8)	4.2 (0.4)	1.8 (0.7)	2.3 (0.9)	53.0 (6.5)	91.3
B	90	75.6	19.3 (1.5)	4.4 (0.9)	4.4 (0.4)	2.1 (0.7)	2.3 (0.8)	57.6 (6.2)	7.8 (1.2)	56.2 (9.2)	27	74.1	22.9 (1.3)	4.0 (0.8)	4.1 (0.4)	1.9 (0.7)	2.3 (0.8)	52.0 (7.9)	92.3
C	122	77.9	18.5 (0.9)	4.2 (0.9)	4.2 (0.4)	2.0 (0.6)	2.3 (0.8)	54.9 (6.2)	6.8 (0.8)	58.7 (5.8)	58	75.9	22.5 (1.6)	4.1 (0.8)	4.2 (0.4)	1.9 (0.7)	2.3 (0.8)	52.6 (5.4)	82.5
Admission group																			
Top-pu-GPA	49	59.2	18.1 (0.6)	4.3 (0.9)	4.2 (0.6)	2.1 (0.8)	2.1 (1.0)	52.5 (7.5)	7.8 (0.9)	58.5 (8.6)	27	63.0	21.7 (1.0)	3.7 (0.9)	4.0 (0.4)	2.2 (0.6)	1.8 (0.6)	47.9 (7.9)	96.0
Lottery	41	78.0	20.1 (2.7)	4.3 (0.9)	4.2 (0.5)	2.1 (0.7)	2.1 (0.9)	54.9 (6.9)	6.9 (1.3)	51.2 (15.1)	116	72.4	22.9 (2.2)	3.9 (0.9)	4.1 (0.5)	1.9 (0.7)	2.2 (0.9)	51.7 (6.9)	76.1
Selection	287	76.7	18.7 (1.2)	4.3 (0.8)	4.3 (0.4)	2.0 (0.6)	2.3 (0.8)	55.6 (6.8)	7.0 (1.1)	56.9 (8.7)	108	76.9	22.7 (1.7)	4.1 (0.8)	4.2 (0.4)	1.9 (0.7)	2.3 (0.6)	52.5 (6.3)	86.8

ENG engagement, *AM* autonomous motivation, *CM* controlled motivation, *RAM* relative autonomous motivation, *SMMS* strength of motivation for medical school

Table 3 Results regression analyses Year-1

Year-1	Engagement B	AM B	CM B	RAM B	Strength of motivation B	Course credits B	GPA B
No participation in selection[a]	(0.122) 0.064[1–6]	(−0.094) −0.105[5;7]	(−0.042) −0.014[7,8]	(−0.022) 0.052[3;5;7]	(−2.149) −0.458[1;7;9;10]	(−5.535**) −1.528[6;9;11]	(−0.594*) 0.046[3;6;9;12]
% of variance explained by final model	38%	47%	3%	23%	37%	15%	22%
Procedure[bd]							
A (VUmc)	(0.175) 0.030[2;4–6;12]	(0.048) 0.009[5;7]	(−0.044) −0.017[7]	(0.083) 0.027[5;7;12]	(−0.652) −1.060[1;7]	(−3.858***) **−3.404**[6]	(−0.339**) −0.253[6]
B (AMC)	(0.170) −0.086[2;4–6;12]	(0.143**) 0.033[5;7]	(0.110) 0.146[7]	(0.012) −0.048[5;7;12]	(2.791*) **1.710***[1;7]	(−2.521) −2.262[6]	(0.947***) **0.995*****[6]
% of variance explained by final model	35%	46%	3%	23%	38%	11%	36%
Admission group[cd]							
Top-pu-GPA	(−0.071) 0.114[1;45;13]	(−0.123) −0.062[5;7]	(0.0.58) 0.058	(−0.161) −0.070[5;7]	(−3.053**) **−2.581*****[1;7;9]	(1.553) 3.022[1;9;11]	(**0.738***) **0.526***[9]
Lottery	(0.017) 0.084[1;45;13]	(−0.084) −0.086[5;7]	(0.136) 0.136	(−0.214) −0.215[5;7]	(−0.664) 0.491[1;7;9]	(−5.726***) −4.004[1;9;11]	(−0.105) 0.124[9]
% of variance explained by final model	42%	45%	0%	23%	33%	9%	6%

In parentheses: unadjusted for confounders; without parentheses: final model

Confounders included in the final model were AM[1], RAM[2], being first generation student[3], living situation[4], strength of motivation[5], pu-GPA[6], engagement[7], area of growing up[8], university[9], ethnicity[10], gender[11], age[12] and CM[13]

Bold indicate significant values in the final model

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

[a] Reference group: participation in selection

[b] Reference group: UMCG

[c] Reference group: selection

[d] We applied the Bonferroni procedure for multiple comparison correction

Table 4 Results regression analyses Year-4

Year-4	Engagement B	AM B	CM B	RAM B	Strength of motivation B	Good clerkship performance ODDS
No participation in selection[a]	(−0.415***) **−0.317***[1;2]**	(−0.107) 0.015[2−5]	(0.131) 0.131	(−0.254) −0.176[2;3]	(−0.662) 0.470[1;3;6]	(0.347***) **0.347***
% of variance explained by final model	40%	31%	1%	14%	32%	4%
Procedure[bc]						
A (VUmc)	(0.186) 0.244[1;2;7;8]	(−0.034) −0.056[2;3;7−9]	(−0.044) −0.045[9]	(0.032) −0.002[2;3;9]	(0.461) 0.124[1;3;6−8;10;11]	(0.448) 0.448
B (AMC)	(−0.053) 0.064[1;2;7;8]	(−0.058) −0.082[2;3;7−9]	(0.000) 0.055[9]	(0.003) −0.096[2;3;9]	(−0.548) 1.502[1;3;6−8;10;11]	(0.392) 0.392
% of variance explained by final model	43%	41%	4%	18%	52%	52%
Admission group[£c]						
Top-pu-GPA	(−0.370*) 0.005[2;8;12;13]	(−0.182) 0.041[2−5;9]	(0.317) 0.317	(−0.468***) −0.325[2;3]	(−4.620***) −2.124*[1;3;4;6]	(1.096) **1.218***[8;12]
Lottery	(−0.224*) −0.136[2;8;12;13]	(−0.064) 0.019[2−5;9]	(0.064) 0.064	(−0.140) −0.077[2;3]	(−0.831) 0.017[1;3;4;6]	(0.899**) 0.918[8;12]
% of variance explained by final model	29%	42%	2%	39%	34%	8%

In parentheses: unadjusted for confounders; without parentheses: final model

Confounders included in the final model were AM[1], strength of motivation[2], engagement[3], gender[4], CM[5], age[6], ethnicity[7], being first generation student[8], area of growing up[9], living situation[10], pu-GPA[11], RAM[12] and university[13]

Bold indicate significant values in the final model

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

[a] Reference group: participation in selection

[b] Reference group: UMCG

[c] We applied the Bonferroni procedure for multiple comparison correction

[£] Reference group: selection

only in Year-1. Selected students did not show better quality of motivation and engagement than the other students.

Participation in selection

Students who had participated in selection were more likely to show good performance during their clerkships (odds ratio 2.883, $p < 0.01$) and reported significantly higher engagement in Year-4 (B = 0.317, $p < 0.05$) than student who had not participated. Analyses showed no significant associations with performance and engagement in Year-1, AM, CM, RAM and strength of motivation. Our hypothesis that students who participated in selection would outperform others and show better motivation and engagement was not supported for Year-1, because there were no differences in this regard. Our hypothesis was partly supported for Year-4 students.

Type of selection procedure

Year-1

Analysis showed significant associations between type of selection procedure and performance and strength of motivation in Year-1. Selection C was associated with more course credits than selection procedure A (B = 3.404, $p < 0.05$). Procedure B was associated with higher GPAs than Procedures A (B = 1.248, $p < 0.01$) and C (B = 0.995, $p < 0.01$). In addition, Procedure B was associated with higher strength of motivation in Year-1 than Procedures A (B = 2.770, $p < 0.01$) and C (B = 1.170, $p < 0.1$). Analyses showed no significant associations with performance in Year-4, AM, CM, RAM, engagement and strength of motivation in Year-4. Our hypothesis that students admitted through the three different selection procedures would not show differences was supported for Year-4 and only partly supported for Year-1. In Year-1, type of motivation and engagement were similar among students selected through the different procedures. Procedure B was associated with higher GPAs and strength of motivation than Procedures A and C, and Procedure C was associated with more course credits.

Discussion

Building on previous literature, this multi-site study investigated the added value of selection compared to a weighted lottery procedure by focusing on student performance, motivation and engagement in both pre-clinical and clinical phase of the medical study. Findings with regard to the different admission groups confirmed that students who excel in pre-university education perform better in the pre-clinical and clinical phases of medical study as well, despite showing lower strength of motivation than selected students. This was not surprising as previous performance is the best predictor of future performance (Benbassat and Baumal, 2007; Hulsman et al. 2007; Patterson et al. 2016; Salvatori 2001; Siu and Reiter 2009). Lower strength of motivation among top-GPA students has been reported before (Hulsman et al. 2007; Kusurkar et al. 2010; Wouters et al. 2016) and may be explained by the fact that these students gain direct admission to the medical study. Without the need for participating in a selection or weighted lottery procedure, these students may be stimulated less to think about their study choice, an activity which can

help in making an informed and motivated study choice (Wouters et al. 2014). In our study, selected students did not outperform lottery-admitted students or report better quality of motivation and engagement in the pre-clinical and clinical phase of the study, which supports the notion that selection may have little added value over a weighted lottery procedure. However, findings differ greatly across studies, suggesting that context is an important factor. The increased proportion of places allocated through selection, for example, may stimulate a broader range of students to apply for selection, reducing the performance gap between selected and lottery-admitted students. Moreover, selection tools are used in different ways in different contexts, which complicates replication of studies and generalization of findings (Edwards et al. 2013). Furthermore, differences are usually small and do not always reach significance. For motivation and engagement this may be due to the restricted range, which means that students scored at the top end for engagement and autonomous motivation and at the bottom end for controlled motivation. The mean engagement score in our sample (M = 4.17), for example, clearly exceeds the score of the norm group of social sciences students (M = 3.18) (Schaufeli and Bakker 2003). In sum, top-pu-GPA students outperformed the other students, while few differences were found between selected and lottery-admitted students. This highlights the challenge that selection committees are confronted with, namely selecting the best candidates from a pool of seemingly equally suitable candidates. Future research should reveal whether students' performance, motivation and engagement develop differently throughout medical study. A next important step in selection research is to follow up on the various groups of students after graduation. We plan to conduct longitudinal research to study this.

Our hypothesis that students who had participated would outperform and show higher AM and engagement than students who had not participated was only partly supported. As no differences in motivation were found, the assumption that better motivation among selection participants would explain their better performance (Schripsema et al. 2014) was not supported. Among students in the clinical phase, participation in selection was related with better clerkship performance and engagement. Students who previously chose to participate in a selection procedure, for which coping with stress and being able to combine studies with other activities are important, may become energized by and be able to cope better with the pressure of clerkships. Students who had participated in selection have been found to be more emotionally stable and conscientious than students who did not (Schripsema et al. 2016). The group of Year-1 students that had not participated in selection was rather small; this might explain why these differences did not reach significance.

Based on the hypothesis that the presence of a selection procedure may be more important than the type of procedure used, we did not expect to find differences in performance, motivation and engagement. Findings among the students in the clinical phase supported this. This must be interpreted with caution, however. Relatively small group sizes might have resulted in insufficient power to detect smaller effects. Among the students in the pre-clinical phase, we found some differences, mainly related to performance. Of course, the medical school context, as a whole, should be considered when interpreting these results (Edwards et al. 2013). While the three medical schools train their students to meet the same end terms, differences in curricular structures and assessment and grading programs may have influenced the study results. Indeed, the factor 'university' appeared to be a confounder in some of the other analyses, but could not be controlled for in the comparisons between the selection procedures. Differences with regard to cognitive performance in the medical study may be related to the weightings of cognitive assessments in the selection procedures. The findings seem to suggest that potential differences between

the students selected at the three medical schools fade over the course of medical study, but a longitudinal study design is necessary to confirm this. Another explanation may be that the characteristics of the three medical schools (selection procedure, curriculum and location) appeal to different types of students. Two of the three medical schools in our study are located in the same city. We have examined different types of students' reasons for applying to a certain medical school in a separate paper (Wouters et al. submitted). In sum, few differences were found between students admitted through the different selection procedures. Because the differences mainly concerned performance outcomes, they may have been strongly influenced by differences in the assessments and grading cultures of the different institutions. Further research should determine the relative influence of curriculum characteristics on performance differences.

Some of the outcome measures in the study were interrelated. For example, autonomous motivation showed a low positive correlation with course credits. Low negative correlations were found between clerkship performance and controlled motivation, while low positive correlations were found between clerkship performance and engagement and relative autonomous motivation. This is in line with the previous research on motivation and engagement reporting positive correlations with performance. Moreover, some significant differences found in the unadjusted models disappeared when confounders were included in the final model. For performance outcomes, age, gender, pu-GPA and university mainly caused this, but sometimes also socioeconomic factors, such as being a first generation student or ethnic background. Further research should determine the influence of socioeconomic factors on student performance. The final models explained up to 52% of the variance in the outcome measures. For some measures, e.g. controlled motivation and course credits, the models explained little variance, which indicates the need for more research on these outcomes.

Limitations

Possible limitations include selection bias and response bias. While we included the three medical schools for methodological and practical reasons, the findings may not be generalizable to other medical schools. In addition, administration of a web-based survey enabled us to approach all students from the proposed cohorts, decreasing the influence of selection bias at the student level, but also may have resulted in a lower response rate. Nevertheless, a response rate of 35% can be considered good in current times in which students receive many evaluation forms and junk mail (Sax et al. 2003). Female students were slightly overrepresented in our study. We included gender as a confounder in the analyses whenever necessary. A response bias is likely because we do not know how non-responders would have answered the motivation and engagement questions. It is reasonable to assume that non-responders have lower motivation and engagement. Some groups in our sample were relatively small. We have taken this into account in the interpretation of the findings. The top-pu-GPA group is consistently small because only 4% of all pre-university graduates in the Netherlands achieve this. A further limitation is that the clerkship grade in the first year of clinical rotations may not be a true reflection of how the students will perform as doctors in their actual practice owing to the little autonomy they have. Furthermore, the grades may reflect students' ability to cope with the transition from theory to practice, rather than their clinical skills. Future research on performance in later stages of medical education and specialty training could provide more insight in more clinically relevant outcomes of selection.

Conclusion

Top performing students in pre-university education perform best in the medical study. A selection, which is usually costly, seems to be of little additional value compared to a weighted lottery procedure, especially when a large proportion of students is admitted through selection. The results suggest that the type of selection procedure may make little difference. Differences are small due to good overall performance, motivation and engagement levels.

Author contributions All authors made substantial contributions to the conception, analysis and/or design of the work. AW led the data acquisition and analysis and prepared the first draft of this paper. NRS and GWGS were responsible for the retrieval of data from UMCG and AMC, respectively. All authors contributed to the critical revision of the document through several iterations. All authors approved the final manuscript for publication.

Compliance with ethical standards

Conflict of interest The authors report no conflict of interests.

Ethical approval Informed consent was obtained from all participants. The data were anonymized before analyses. The study was approved by the Ethical Review Board of the Netherlands Association for Medical Education (NVMO-ERB, dossier number 266).

References

Artino, A. R., La Rochelle, J. S., & Durning, S. J. (2010). Second-year medical students's motivational beliefs, emotions, and achievement. *Medical Education, 44*(12), 1203–1212. doi:10.1111/j.1365-2923. 2010.03712.x.

Benbassat, J., & Baumal, R. (2007). Uncertainties in the selection of applicants for medical school. *Advances in Health Sciences Education, 12,* 509–521. doi:10.1007/s10459-007-9076-0.

Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in Higher Education, 47,* 1–32. doi:10.1007/s11162-005-8150-9.

Casuso-Holgado, M. J., Cuesta-Vargas, A. I., Moreno-Morales, N., Labajos-Manzanares, M. T., Barón-López, F. J., & Vega-Cuesta, M. (2013). The association between academic engagement and achievement in health sciences students. *BMC Medical Education, 13,* 1–7. doi:10.1186/1472-6920-13-33.

Cleland, J., Dowell, J., McLachlan, J., Nicholson, S., & Patterson, F. (2012). Identifying best practice in the selection of medical students. GMC Research Report

Daelmans, H. E., Mak-van der Vossen, M., Croiset, G., & Kusurkar, R. A. (2016). What difficulties do faculty members face when conducting workplace-based assessments in undergraduate clerkships? *International Journal of Medical Education, 7,* 19–24.

De Visser, M., Fluit, C., Fransen, J., Latijnhouwers, M., Cohen-Schotanus, J., & Laan, R. (2016). The effect of curriculum sample selection for medical school. *Advances in Health Sciences Education.* doi:10.1007/s10459-016-9681-x.

Deci, E. L., & Ryan, R. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum.

Edwards, D., Friedman, T., & Pearce, J. (2013). Same admissions tools, different outcomes: A critical perspective on predictive validity in three undergraduate medical schools. *BMC Medical Education, 13,* 1. doi:10.1186/1472-6920-13-173.

Hulsman, R. L., Van Der Ende, J. S. J., Oort, F. J., Michels, R. P. J., Casteelen, G., & Griffioen, F. M. M. (2007). Effectiveness of selection in medical school admissions: Evaluation of the outcomes among freshmen. *Medical Education, 41,* 369–377. doi:10.1111/j.1365-2929.2007.02708.x.

Kusurkar, R. A., Croiset, G., Kruitwagen, C., & Cate, T. J. (2011a). Validity evidence for the measurement of the strength of motivation for medical school. *Advances in Health Sciences Education, 16,* 183–195. doi:10.1007/s10459-010-9253-4.

Kusurkar, R. A., Kruitwagen, C., Ten Cate, T. J., & Croiset, G. (2010). Effects of age, gender and educational background on strength of motivation for medical school. *Advances in Health Sciences Education, 15,* 303–313. doi:10.1007/s10459-009-9198-7.

Kusurkar, R. A., Ten Cate, T. J., Van Asperen, M., & Croiset, G. (2011b). Motivation as an independent and a dependent variable in medical education: A review of the literature. *Medical Teacher, 33,* e242–e262. doi:10.3109/0142159X.2011.558539.

Kusurkar, R. A., Ten Cate, T. J., Vos, C. M. P., Westers, P., & Croiset, G. (2013). How motivation affects academic performance: A structural equation modelling analysis. *Advances in Health Sciences Education, 18,* 57–69. doi:10.1007/s10459-012-9354-3.

Leibach, G., & Stern, M. (2013, December 27). Critical synthesis package: Strength of motivation for medical school-revised (SMMS-R) questionnaire. MedEdPORTAL Publications, 9, Publication ID 9645. Retrieved from https://www.mededportal.org/publication/9645

Lucieer, S. M., Stegers-Jager, K. M., Rikers, R. M., & Themmen, A. P. (2015). Non-cognitive selected students do not outperform lottery-admitted students in the pre-clinical stage of medical school. *Advances in Health Sciences Education.* doi:10.1007/s10459-015-9610-4.

Moulaert, V., Verwijnen, M. G. M., Rikers, R., & Scherpbier, A. J. J. J. A. (2004). The effects of deliberate practice in undergraduate medical education. *Medical Education, 38,* 1044–1052. doi:10.1111/j.1365-2929.2004.01954.x.

Nieuwhof, M. G., Ten Cate, T. J., Oosterveld, P., & Soethout, M. (2004). Measuring strength of motivation for medical school. *Medical Education Online, 9,* 1–7.

Patterson, F., Knight, A., Dowell, J., Nicholson, S., Cousans, F., & Cleland, J. (2016). How effective are selection methods in medical education? A systematic review. *Medical Education, 50,* 36–60. doi:10.1111/medu.12817.

Prins, J. T., Van Der Heijden, F. M. M. A., Hoekstra-Weebers, J. E. H. M., Bakker, A. B., Van de Wiel, H. B. M., Jacobs, B., et al. (2009). Burnout, engagement and resident physicians' self-reported errors. *Psychology, Health and Medicine, 14,* 654–666. doi:10.1080/13548500903311554.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25,* 54–67. doi:10.1006/ceps.1999.1020.

Salvatori, P. (2001). Reliability and validity of admissions tools used to select students for the health professions. *Advances in Health Sciences Education, 6,* 159–175. doi:10.1023/A%3A1011489618208.

Sax, L. J., Gilmartin, S. K., & Bryant, A. N. (2003). Assessing response rates and nonresponse bias in web and paper surveys. *Research in Higher Education, 44*(4), 409–432. doi:10.1023/A:1024232915870.

Schaufeli, W. B., & Bakker, A. B. (2003). UWES ÇôUtrecht work engagement scale: Test manual. Unpublished manuscript, Department of Psychology, Utrecht University.

Schaufeli, W. B., Martinez, I. M., Pinto, A. M., Salanova, M., & Bakker, A. B. (2002a). Burnout and engagement in university students a cross-national study. *Journal of Cross-Cultural Psychology, 33*(5), 464–481.

Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002b). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness Studies, 3,* 71–92. doi:10.1023/A:1015630930326.

Schripsema, N. R., Trigt, A. M., Borleffs, J. C., & Cohen-Schotanus, J. (2014). Selection and study performance: Comparing three admission processes within one medical school. *Medical Education, 48,* 1201–1210. doi:10.1111/medu.12537.

Schripsema, N. R., van Trigt, A. M., van der Wal, M. A., & Cohen-Schotanus, J. (2016). How different medical school selection processes call upon different personality characteristics. *PLoS ONE, 11,* e0150645. doi:10.1371/journal.pone.0150645.

Siu, E., & Reiter, H. (2009). Overview: What's worked and what hasn't as a guide towards predictive admissions tool development. *Advances in Health Sciences Education, 14,* 759–775. doi:10.1007/s10459-009-9160-8.

Sobral, D. T. (2004). What kind of motivation drives medical students' learning quests? *Medical Education, 38,* 950–957. doi:10.1111/j.1365-2929.2004.01913.x.

Stegers-Jager, K. M., Cohen-Schotanus, J., & Themmen, A. P. N. (2012). Motivation, learning strategies, participation and medical school performance. *Medical Education, 46,* 678–688. doi:10.1111/j.1365-2923.2012.04284.x.

Stegers-Jager, K. M., Themmen, A. P. N., Cohen-Schotanus, J., & Steyerberg, E. W. (2015). Predicting performance: Relative importance of students' background and past performance. *Medical Education, 49,* 933–945. doi:10.1111/medu.12779.

Strauser, D. R., O'Sullivan, D., & Wong, A. W. (2012). Work personality, work engagement, and academic effort in a group of college students. *Journal of Employment Counseling, 49,* 50–61. doi:10.1002/j.2161-1920.2012.00006.x.

Svanum, S., & Bigatti, S. M. (2009). Academic course engagement during one semester forecasts college success: Engaged students are more likely to earn a degree, do it faster, and do it better. *Journal of College Student Development, 50,* 120–132. doi:10.1353/csd.0.0055.

Ten Cate, T. J. (2007). Medical education in the Netherlands. *Medical Teacher, 29,* 752–757. doi:10.1080/01421590701724741.

Turner, R., & Nicholson, S. (2011). Reasons selectors give for accepting and rejecting medical applicants before interview. *Medical Education, 45,* 298–307. doi:10.1111/j.1365-2923.2010.03874.x.

Twisk, J. W. (2006). *Applied multilevel analysis: A practical guide for medical researchers.* Cambridge: Cambridge University Press.

Urlings-Strop, L. C., Stegers-Jager, K. M., Stijnen, T., & Themmen, A. P. N. (2013). Academic and non-academic selection criteria in predicting medical school performance. *Medical Teacher.* doi:10.3109/0142159X.2013.774333.

Urlings-Strop, L. C., Stijnen, T., Themmen, A. P. N., & Splinter, T. A. W. (2009). Selection of medical students: A controlled experiment. *Medical Education, 43,* 175–183. doi:10.1111/j.1365-2923.2008.03267.x.

Urlings-Strop, L. C., Themmen, A. P. N., Stijnen, T., & Splinter, T. A. W. (2011). Selected medical students achieve better than lottery-admitted students during clerkships. *Medical Education, 45,* 1032–1040. doi:10.1111/j.1365-2923.2011.04031.x.

Van Herwaarden, C. L. A., Laan, R. F. J. M., & Leunissen, R. R. M. (2009). *The 2009 framework for undergraduate medical education in the Netherlands.* Utrecht: Dutch Federation of University Medical Centres.

Vansteenkiste, M., Sierens, E., Soenens, B., Luyckx, K., & Lens, W. (2009). Motivational profiles from a self-determination perspective: The quality of motivation matters. *Journal of Educational Psychology, 101,* 671–688. doi:10.1037/a0015083.

Vansteenkiste, M., Zhou, M., Lens, W., & Soenens, B. (2005). Experiences of autonomy and control among Chinese learners: Vitalizing or immobilizing? *Journal of Educational Psychology, 97,* 468–483. doi:10.1037/0022-0663.97.3.468.

Williams, G. C., Saizow, R. B., & Ryan, R. M. (1999). The importance of self-determination theory for medical education. *Academic Medicine, 74,* 992–995.

Wouters, A., Bakker, A. H., van Wijk, I. J., Croiset, G., & Kusurkar, R. A. (2014). A qualitative analysis of statements on motivation of applicants for medical school. *BMC Medical Education, 14*(1), 200. doi:10.1186/1472-6920/14/200.

Wouters, A., Croiset, G., Galindo-Garre, F., & Kusurkar, R. A. (2016). Motivation of medical students: Selection by motivation or motivation by selection. *BMC Medical Education, 16,* 1. doi:10.1186/s12909-016-0560-1.

Wouters, A., Croiset, G., Schripsema, N. R., Cohen-Schotanus, J., Spaai, G. W. G., Hulsman, R. L., & Kusurkar, R. A. Medical schools' and students' approaches to selection: Conflict of interest? **(submitted)**