Applicant perceptions of selection methods for health professions education: Rationales and subgroup differences

Suzanne Fikrat-Wevers[1] | **Karen Stegers-Jager**[1] | **Marleen Groenier**[2] | **Andries Koster**[3] | **Jan Hindrik Ravesloot**[4] | **Renske Van Gestel**[3] | **Anouk Wouters**[5,6] | **Walter van den Broek**[1] | **Andrea Woltman**[1]

[1]Institute of Medical Education Research Rotterdam, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

[2]Technical Medical Centre, Technical Medicine, University of Twente, Enschede, The Netherlands

[3]Department of Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands

[4]Department of Medical Biology, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

[5]Faculty of Medicine VU, Amsterdam UMC location Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

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

Correspondence
Andrea Woltman, Institute of Medical Education Research Rotterdam, Erasmus MC, University Medical Center Rotterdam, Room AE-231, PO Box 2040, 3000 CA Rotterdam, The Netherlands.
Email: a.woltman@erasmusmc.nl

Funding information
This study is supported by Nationaal Regieorgaan Onderwijsonderzoek (40.5.18650.007).

Abstract

Context: Applicant perceptions of selection methods can affect motivation, performance and withdrawal and may therefore be of relevance in the context of widening access. However, it is unknown how applicant subgroups perceive different selection methods.

Objectives: Using organisational justice theory, the present multi-site study examined applicant perceptions of various selection methods, rationales behind perceptions and subgroup differences.

Methods: Applicants to five Dutch undergraduate health professions programmes ($N = 704$) completed an online survey including demographics and a questionnaire on applicant perceptions applied to 11 commonly used selection methods. Applicants rated general favourability and justice dimensions (7-point Likert scale) and could add comments for each method.

Results: Descriptive statistics revealed a preference for selection methods on which applicants feel more 'in control': General favourability ratings were highest for curriculum-sampling tests (mean [M] = 5.32) and skills tests ($M = 5.13$), while weighted lottery ($M = 3.05$) and unweighted lottery ($M = 2.97$) were perceived least favourable. Additionally, applicants preferred to distinguish themselves on methods that assess attributes beyond cognitive abilities. Qualitative content analysis of comments revealed several conflicting preferences, including a desire for multiple selection methods versus concerns of experiencing too much stress. Results from a linear mixed model of general favourability indicated some small subgroup differences in

perceptions (based on gender, migration background, prior education and parental education), but practical meaning of these differences was negligible. Nevertheless, concerns were expressed that certain selection methods can hinder equitable admission due to inequal access to resources.

Conclusions: Our findings illustrate that applicants desire to demonstrate a variety of attributes on a combination of selection tools, but also observe that this can result in multiple drawbacks. The present study can help programmes in deciding which selection methods to include, which more negatively perceived methods should be better justified to applicants, and how to adapt methods to meet applicants' needs.

1 | INTRODUCTION

Due to the high stakes involved, selection into undergraduate health professions education (HPE) is a topic of public debate.[1] Nevertheless, most research into different selection methods has focused on predictive validity and effects on student diversity, while little attention has been paid to applicant perceptions.[2] This is surprising, as it has been argued that applicant perceptions are relevant within the context of widening access (WA).[1] Although the underrepresentation of lower socio-economic and ethnic minority groups is a complex and multicausal problem, these groups are thought to be particularly vulnerable to the consequences of negative perceptions regarding selection.[1,3] However, thus far, no research has compared perceptions between different subgroups of applicants to investigate whether this is really the case. In the present study, we addressed this knowledge gap by investigating applicant perceptions of different selection methods in undergraduate HPE and how these perceptions are related to applicants' background characteristics.

Understanding applicant perceptions of selection methods is important for multiple reasons. First, applicant perceptions are a component of political validity, which is used as a source of evidence in broader validity frameworks.[4] Political validity is defined as the extent to which stakeholders consider a method as appropriate and acceptable for use in selection.[4] In practice, political validity is taken into consideration when designing selection procedures, which is exemplified by the fact that selection methods with proven limited predictive validity, such as traditional interviews, are still commonly used.[1] Second, perceptions of applicants are of particular interest, as they can have numerous practical consequences, including effects on construct validity of selection methods, and on applicant motivation, performance and withdrawal.[2,5-7] Noteworthy is that applicants' understanding of the expectations of selection committees can shape the way they prepare for selection in order to 'fit' these hidden expectations.[8-10] Moreover, applicant perceptions are relevant with respect to WA.[1] Previous research suggests that prospective students with lower socio-economic and ethnic minority backgrounds more often demonstrate negative perceptions of selection and are less confident in their ability to be successfully admitted.[3,11] These subgroups of applicants are shown to have poorer selection outcomes in general,[12,13] which may be partially explained by poorer motivation due to more negative perceptions.[14]

Applicant perceptions of selection are commonly described using organisational justice theory,[15] which distinguishes between distributive and procedural justice. Distributive justice describes the perceived fairness of the distribution of the outcome of the selection procedures, while procedural justice refers to the perceived fairness of the selection procedure and methods that are used to generate this outcome.[15] Procedural justice consists of three components: the formal characteristics of procedures, explanation of procedures and decision-making and interpersonal treatment. According to organisational justice theory, the three procedural justice components affect perceptions of different aspects of selection procedures. The present study focuses on the procedural justice component formal characteristics of selection, because according to Gilliland,[15] perceptions of specific selection methods—the main interest of the present study—are mostly affected by this component. Examples of such formal characteristics include relevance and perceived validity of selection methods. The other two components, explanation of procedures and decision-making and interpersonal treatment, influence perceptions of the selection policy and the selection personnel, respectively.

Educational institutions use a great variety of methods in their selection procedures. However, knowledge about perceptions of applicants regarding these methods is still scarce.[2] A systematic review concluded that interviews and situational judgement tests (SJTs) are highly supported by applicants, while aptitude tests are perceived as less acceptable.[1] Although this review included a large number of studies (71), most were conducted in single institutions after exposure to one specific selection method. Because previous research indicates that the selection procedure plays a role in study choice,[2,16] results of such studies are only limitedly comparable and generalisable. Additionally, while Kelly et al.[1] suggested that certain selection methods are specifically perceived as barriers by underrepresented and minority groups, none of the studies under review directly compared perceptions of different applicant subgroups.

Thus far, only few studies have directly investigated subgroup differences in applicant perceptions. To our knowledge, only one single-site study was executed in the field of HPE. This research focused on SJTs and concluded that ethnic minority students and first-generation university students had other preferences in SJT format compared with their traditional counterparts.[14] Another single-site study that was conducted at an undergraduate psychology programme included a wide range of selection methods but only focused on gender as a

background variable.[2] Low favourability ratings were found for lottery and high school grades, while tests and interviews were perceived as highly acceptable, but the rationales behind the provided ratings were not explored. The authors found that only motivation questionnaires were rated more positively by women than by men, while perceptions of other selection methods under research did not differ by gender.

Thus, there is a gap in the applicant perceptions literature with respect to research in multi-institutional settings, that includes a variety of methods and that includes numerous relevant background characteristics. In the present study, we investigated how applicants perceive different selection methods for admission into undergraduate HPE and the rationales behind their perceptions. Additionally, we assessed whether these perceptions differ across subgroups of applicants based on background variables that are associated with selection outcomes: applicants' gender, migration background (as an indicator of ethnicity), parental education (as an indicator of SES) and prior education.[12,13,17,18]

2 | METHOD

2.1 | Setting and procedure

The present study was conducted at five undergraduate HPE programmes in the Netherlands, including three medical programmes, one technical-medical programme and one pharmacy programme. All five programmes had different self-designed selection procedures, composed of at least two selection methods.

In the Netherlands, admission requirements of different types of undergraduate HPE programmes are comparable. To be eligible, all applicants need to meet the same stringent requirements regarding subjects taken (e.g. physics, chemistry and biology) and educational level (graduation level of pre-university education). Although applicants can apply from different educational routes, they all need to provide proof that they meet aforementioned requirements. Consequently, the applicant pools are relatively homogeneous; students who apply to a university-level undergraduate HPE programme are already strongly preselected based on academic skills due to highly selective secondary education.[19] There are no admission requirements, however, with respect to pu-GPA. When applicants apply to their programme of choice, they apply to one specific institution. For each institution, there is a predetermined fixed number of places. By law, institutions are required to include at least two selection criteria, which are assessed using selection methods—the individual tools that can be implemented as part of a selection procedure. However, there are no additional requirements with respect to, for instance, the contents and quality of the selection methods, and in which way they would be combined. Consequently, there is great variety in the selection procedures that programmes employ, both between and within different types of HPE programmes at different institutions. Currently, a bill is pending that should allow programmes to also include (weighted and unweighted) lottery as a selection method as part of their selection procedure.

All applicants engaged in the selection procedures for entry in September 2020 ($N = 3280$) were invited to participate in an online survey on applicant perceptions. Additionally, applicants were asked to complete a demographics questionnaire either during an on-site testing day or via e-mail.

Informed consent was obtained from all participants. Applicants were informed that participation was voluntary and would not influence their selection outcomes. No incentives were provided. The Medical Ethical Review Committee of Erasmus MC declared the study exempt from review.

2.2 | Measures

Applicant perceptions were measured using a previously reported questionnaire including seven items based on organisational justice theory.[14] Each item was judged using a 7-point Likert scale. In accordance with previous studies,[2,14] general favourability was assessed using two items: perceived predictive validity and perceived fairness. The other five items reflected the following procedural justice dimensions: face validity, applicant differentiation, study relatedness, chance to perform and ease of cheating. The specific items (i.e. definitions provided to respondents) and rating scales of each justice dimension can be consulted in Table S1. Respondents were asked to rate the applicant perception items separately for 11 commonly used selection methods: cognitive capacity test, curriculum-sampling test, curriculum vitae (CV), interview, motivation letter, personality questionnaire, pre-university grade-point average (pu-GPA), SJT, skills test, unweighted lottery and weighted lottery. For unweighted and weighted lottery, only general favourability and face validity were assessed, as the other items were considered irrelevant for those methods. A short description of each selection method was provided (Table S2), along with example items for the SJT and personality questionnaire. For each method (including lottery), applicants were also asked to answer the following open-ended question: 'Do you have any remarks about using [insert method] as a selection method?' The selection methods were presented in a random order to mitigate order effects.

The following variables were included in the demographics questionnaire: gender, migration background, prior education and parental education. Gender diversity was acknowledged in the present study, and applicants had the option to choose between three categories: 'man', 'woman' and 'other, namely [free text box]'. Migration background was defined in accordance with Statistics Netherlands (CBS). Applicants were categorised in three groups: (i) no migration background when both parents were born in the Netherlands; (ii) a Western migration background when at least one parent was born in Europe (excluding the Netherlands and Turkey), North America, Oceania, Japan or Indonesia; and (iii) a non-Western migration background when at least one parent was born in Africa, Asia (excluding Japan and Indonesia), Latin America or Turkey. With respect to prior education, we distinguished between standard Dutch pre-university education, university and other forms of prior education (e.g. higher vocational

education, foreign education). Finally, for parental education, applicants were categorised as first-generation university applicants when none of their parents had attended higher education (university or higher vocational education).

2.3 | Analyses

In order to investigate applicant perceptions of different selection methods, we used descriptive statistics. After checking for internal consistency, general favourability of each method for each respondent was calculated as the mean score on the two general favourability items. For general favourability and the other items, the mean score and 95% confidence interval for each selection method were calculated. Additionally, we calculated correlations between scores on general favourability and the other items for each method to examine relationships between the procedural justice dimensions.

To study whether applicant perceptions differed for subgroups of applicants, we used a linear mixed model. The mixed model compensated for partially missing data and allowed us to control for individual and context-related differences. The dependent variable in this model was the mean general favourability score. Fixed effects included the selection method, as well as the interaction effects between each selection method and the demographic variables of interest (gender, migration background, prior education and parental education). Random factors included applicant id and the programme to which the applicant applied, using an unstructured covariance structure. The decision on the covariance structure was based on the Akaike information criterion. For this analysis, data of individual applicants were used only when at least 10 of the 11 selection methods were rated on general favourability.

The answers to the open-ended questions were used to get a more in-depth view of the rationales behind applicant perceptions, from a constructivist paradigm. We used qualitative analysis of content, employing a directed approach.[20] With a directed approach, a theory is used as initial guidance for the formation of codes with the goal of validating and extending the theoretical framework. In the present study, theoretical concepts of organisational justice theory formed the foundation for the coding process. We followed the analytical procedure described by Zhang and Wildemuth.[21] The first author (SFW) familiarised herself with the data and developed a coding manual. Higher order codes were predetermined and were based on the theoretical concepts of the questionnaire: face validity, applicant differentiation, study relatedness, chance to perform and ease of cheating. Additionally, other concepts of organisational justice theory that were not covered by the questionnaire we administered were included as higher order codes in the coding manual, including distributive justice, reactions during hiring, reactions after hiring and self-perceptions, as well as other items used in previously reported questionnaires referring to procedural justice: scientific evidence, interpersonal warmth, right to use, invasion of privacy, widespread use, effort expectancy, information known, reconsideration opportunity, feedback, consistency of administration, openness,

treatment and two-way communication.[2,22] Lower order codes were developed during the coding process. Thus, we first identified to which theoretical concept an open answer referred (higher order codes) and subsequently coded the specific argument the applicant provided (lower order codes). We also aimed to employ reflexivity to identify potentially relevant findings that were outside of our theoretical framework, by having critical dialogues throughout all stages of the analyses and being aware for blind spots in the theory. SFW (PhD candidate, background in educational sciences) and the second author (KMSJ; senior researcher, background in educational sciences, experienced in research on diversity) coded two answers for each selection method together and then coded a random sample of 10 answers for each selection method independently. Because there was a high level of agreement, SFW coded the remaining answers, after which KMSJ conducted a consistency check. The consistency check entailed that for each code, it was critically reviewed whether a subsample of the comments did indeed apply to the subject covered by that specific code. SFW drew conclusions from the data that were discussed with the full research team.

3 | RESULTS

3.1 | Participant characteristics

In total, 704 applicants participated in the study (response rate = 21%). Amongst the respondents, 71% identified as woman, and one applicant identified as 'other'. This individual was excluded from the subgroup analyses and therefore only the categories men and women are described in the results. Furthermore, 35% had a migration background (26% non-Western, 9% Western), and 25% were first-generation university applicants. With respect to prior education, 77% applied directly from standard pre-university education, 15% from university and 8% from other forms of prior education. With respect to gender, parental education and prior education, demographic distributions in the present study were comparable with those of other research conducted in the Netherlands, where men and first-generation university applicants are underrepresented.[12,17,18] Applicants with a migration background are also underrepresented in the Netherlands,[12] but composed a relatively large subgroup in the present study, which may be due to the urban setting of most of the included programmes.

3.2 | General favourability

The two general favourability items showed a sufficient to good internal consistency (cognitive capacity test: $\alpha = 0.84$, curriculum-sampling test: $\alpha = 0.74$, CV: $\alpha = 0.83$, interview: $\alpha = 0.77$, motivation letter: $\alpha = 0.79$, personality questionnaire: $\alpha = 0.79$, pu-GPA: $\alpha = 0.85$, SJT: $\alpha = 0.81$, skills test: $\alpha = 0.79$, unweighted lottery: $\alpha = 0.79$, unweighted lottery: $\alpha = 0.86$).

TABLE 1 Descriptive statistics for general favourability for each method

Selection method	N	M	SD	95% CI
Cognitive capacity test	693	4.65	1.32	4.55, 4.74
Curriculum-sampling test	690	5.32	1.08	5.24, 5.40
Curriculum vitae	688	3.95	1.56	3.83, 4.06
Interview	684	5.04	1.23	4.95, 5.13
Motivation letter	694	4.52	1.37	4.42, 4.63
Personality questionnaire	685	3.80	1.40	3.70, 3.91
Pre-university grade-point average	687	3.51	1.51	3.40, 3.62
Situational judgement test	688	4.35	1.28	4.25, 4.44
Skills test	693	5.13	1.36	5.05, 5.22
Unweighted lottery	686	2.97	1.78	2.83, 3.10
Weighted lottery	683	3.05	1.57	2.93, 3.17

Note: N = number of individuals, M = mean, SD = standard deviation, 95% CI = 95% confidence interval.

Applicants provided the highest general favourability ratings for curriculum-sampling tests (mean [M] = 5.32, 95% confidence interval [95% CI; 5.24, 5.40]; Table 1), while weighted lottery (M = 3.05 [2.93, 3.17]) and unweighted lottery (M = 2.97 [2.83, 3.10]) were perceived least favourable.

All correlations between the general favourability score and other applicant perception items were statistically significant ($p < 0.05$), but there was a great range in the strengths of the correlations (see Table 3). The justice dimension that was most strongly related with general favourability was face validity, while the negative correlation between ease of cheating and general favourability was the smallest.

Applicant perceptions appear to differ between different programmes (Table 2). Generally, applicants had more positive perceptions towards selection methods used by the programme they applied to compared with methods that were not used.

3.3 | Justice dimensions

In the next sections, the results for the different justice dimensions will be presented by integrating the quantitative findings of perception scores and the qualitative findings of applicants' reasoning. Only the most salient findings are discussed, and for a complete overview of the quantitative and qualitative findings, we refer to Tables 3 and 4, respectively.

3.3.1 | Face validity

Curriculum-sampling tests and skills tests received the highest ratings on face validity (respectively, M = 5.46, 95% CI [5.37, 5.54], M = 5.27 [5.18, 5.36]), but applicants did not comment on these methods with respect to this dimension. This is illustrative of the observation that, also for other dimensions, applicants mainly focused on negative aspects in their comments. Interviews were also rated highly on the

scale of face validity (M = 5.17 [5.07, 5.27]). Applicants considered attributes such as motivation, personality and social skills important but preferred an oral interview over a written assessment of these attributes. This may also explain the relatively low rating of personality questionnaires (M = 3.87 [3.75, 3.99]). Pu-GPA also received a low rating on face validity (M = 3.70 [3.58, 3.82]), and applicants mentioned that pu-GPA was not of added value as the admission requirements would already provide enough evidence of their capacity to enter the programme. The lowest ratings were provided for weighted and unweighted lottery (respectively, M = 2.99 [2.87, 3.12], M = 2.50 [2.37, 2.63]). Although applicants acknowledged that both types of lottery can reduce stress and pressure for applicants, they communicated a strong desire to distinguish themselves and believed that lotteries would not select the best and most motivated students.

3.3.2 | Applicant differentiation

Skills tests and interviews received relatively high scores on applicant differentiation (respectively, M = 5.25 [5.16, 5.34], M = 5.23 [5.13, 5.33]). These were the only methods for which applicants mentioned that distinguishing skills are assessed and assessors can get an adequate overview of applicants' competencies. For most of the selection methods, however, applicants indicated that the methods distinguish candidates based on other skills than the methods are intended to assess. For instance, according to applicants, motivation letters assess writing skills instead of pure motivation, and selection tests measure test wiseness and preparation time instead of the targeted knowledge or skills. Personality questionnaires and SJTs received the criticism that applicants are distinguished based on their knowledge of social expectations—which they expected to be very basic knowledge—rather than their actual personality or reactions to situations. Pu-GPA received the lowest scores (M = 3.25 [3.13, 3.36]), as applicants mentioned that the same grade is the result of a wide range in motivation, effort and degree of difficulty.

TABLE 2 Descriptive statistics for general favourability by programme in descending order

Selection method	Programme[a]	N	M	SD	95% CI
Cognitive capacity test	A	175	4.51	1.42	4.30, 4.72
	B	167	4.91	1.24	4.72, 5.10
	C	185	4.41	1.39	4.21, 4.61
	D[b]	115	4.97	1.04	4.78, 5.17
	E	51	4.37	1.26	4.02, 4.73
Curriculum-sampling test	A[b]	175	5.24	1.11	5.41, 5.08
	B[b]	164	5.47	1.08	5.31, 5.64
	C[b]	185	5.35	1.01	5.20, 5.50
	D	115	5.03	1.14	4.82, 5.24
	E[b]	51	5.69	0.95	5.42, 5.95
Curriculum vitae	A	175	3.15	1.54	2.92, 3.38
	B[b]	163	4.35	1.41	4.14, 4.57
	C[b]	185	4.76	1.30	4.57, 4.95
	D	114	3.52	1.38	3.26, 3.77
	E	51	3.41	1.54	2.98, 3.84
Interview	A	175	4.94	1.30	4.74, 5.14
	B	158	5.20	1.19	5.01, 5.39
	C	185	5.16	1.27	4.97, 5.34
	D	115	4.95	1.03	4.76, 5.14
	E	51	4.65	1.12	4.33, 4.96
Motivation letter	A	175	4.28	1.33	4.08, 4.48
	B	168	4.78	1.31	4.58, 4.98
	C	185	4.67	1.36	4.47, 4.87
	D	115	4.43	1.29	4.19, 4.67
	E	51	4.23	1.66	3.76, 4.69
Personality questionnaire	A	175	3.49	1.34	3.29, 3.70
	B	160	3.78	1.46	3.56, 4.01
	C	184	4.02	1.41	3.81, 4.22
	D	115	3.79	1.33	3.55, 4.04
	E	51	4.20	1.14	3.88, 4.52
Pre-university grade-point average	A	175	2.85	1.52	2.62, 3.08
	B[b]	161	4.00	1.38	3.79, 4.22
	C[b]	185	3.32	1.39	3.11, 3.52
	D[b]	115	3.86	1.42	3.60, 4.12
	E[b]	51	4.15	1.46	3.73, 4.56
Situational judgement test	A	175	4.39	1.36	4.19, 4.60
	B	162	4.35	1.32	4.14, 4.55
	C	185	4.32	1.24	4.14, 4.50
	D	115	4.17	1.16	3.95, 4.38
	E	51	4.69	1.14	4.37, 5.01
Skills test	A[b]	175	5.29	1.18	5.11, 5.46
	B	167	5.11	1.26	4.92, 5.31
	C	185	5.06	1.17	4.89, 5.23
	D[b]	115	5.14	0.97	4.97, 5.32
	E	51	4.94	1.19	4.61, 5.28

(Continues)

TABLE 2 (Continued)

Selection method	Programme[a]	N	M	SD	95% CI
Unweighted lottery	A	175	3.30	1.82	3.03, 3.58
	B	161	2.65	1.67	2.39, 2.91
	C	185	2.82	1.79	2.56, 3.08
	D	115	2.97	1.77	2.64, 3.30
	E	51	3.35	1.77	2.85, 3.85
Weighted lottery	A	175	2.78	1.50	2.56, 3.00
	B	157	3.25	1.49	3.01, 3.48
	C	185	2.72	1.55	2.49, 2.94
	D	115	3.46	1.58	3.17, 3.75
	E	51	3.64	1.74	3.15, 4.13

Note: N = number of individuals, M = mean, SD = standard deviation, 95% CI = 95% confidence interval.
[a]A = Amsterdam UMC, location AMC; B = Erasmus MC; C = Amsterdam UMC, location VUmc; D = University of Twente; E = Utrecht University.
[b]The method is used by the programme.

3.3.3 | Study relatedness

Compared with general favourability, the mean scores on study relatedness were relatively low for all selection methods (range $M = 2.93$–4.43). A possible explanation is that for each selection method, applicants indicated that only certain relevant skills are assessed, while information on other important attributes is missed. For instance, while cognitive capacity tests can, according to applicants, assess relevant skills such as information processing and problem-solving, social and communicative skills are not taken into account. Thus, applicants believed that none of the methods can fully predict study performance and they preferred a combination of selection methods. Although applicants preferred assessment of a broad range of skills, they also stated that there should be room to develop those skills over the course of the program, introducing a dilemma.

Skills tests and curriculum-sampling tests received the highest scores on study relatedness (respectively, $M = 4.43$ [4.33, 4.53], $M = 4.38$ [4.28, 4.48]). Applicants mentioned that both methods can assess a broad range of skills. Moreover, only for these methods, applicants stated that they can become more acquainted with the programme and its course materials, creating an opportunity for applicants to assess whether they are interested in and able to cope with the programme content. For the lowest scoring methods—pu-GPA and CV (respectively, $M = 3.11$ [3.00, 3.22], $M = 2.93$ [2.82, 3.04])—applicants stated that the assessed skills are too generic and that previous performance is not predictive of future success.

3.3.4 | Chance to perform

The mean scores on chance to perform were also relatively low (range $M = 3.09$–4.90). The qualitative results provide some potentially relevant explanations for this finding. First, applicants mentioned some dilemmas with respect to chance to perform. For instance, they preferred assessment of current knowledge and skills over previous achievements and pointed out drawbacks of snapshot assessments. Second, for each selection method, applicants noticed sources of inequality that can interfere with the chance to perform on a method for certain subgroups. For example, the existence of a network within the medical field was considered useful for building a CV, and applicants with enough financial resources could take commercial coaching for better test results and tutoring for a higher pu-GPA. Skills test and interview scored higher on chance to perform (respectively, $M = 4.90$ [4.80, 5.00], $M = 4.77$ [4.66, 4.89]), and applicants noted that they got the impression that these methods allow for more space to show a broad range of qualities. Pu-GPA received the lowest scores ($M = 3.09$ [2.97, 3.21]), which can be explained by the fact that applicants experienced a lack of control for this selection method, as data had been collected before application.

3.3.5 | Ease of cheating

For the four selection methods scoring highest on ease of cheating—personality questionnaire, motivation letter, SJT and interview (respectively, $M = 6.06$ [5.96, 6.16], $M = 5.72$ [5.62, 5.82], $M = 4.72$ [4.57, 4.86], $M = 4.62$ [4.50, 4.74])—applicants mentioned that social desirability can play a major role. An additional remark about motivation letters is that applicants can easily cheat by letting others write their letter. CV was also considered relatively easy to cheat on ($M = 4.42$ [4.28, 4.56]). Applicants explained that it is easy to forge documents, and only a small sample of CVs is checked on correctness. Applicants did not have any remarks about why it is harder to cheat on the lower scoring selection methods.

3.3.6 | Additional dimensions

Besides the justice dimensions that were covered by the questionnaire, we identified a number of additional dimensions in the answers

TABLE 3 Descriptive statistics and correlation with general favourability for each method on each dimension

Dimension	Selection method	M	SD	95% CI	R[a]
Face validity	Cognitive capacity test	4.66	1.43	4.55, 4.76	0.77
	Curriculum-sampling test	5.46	1.14	5.37, 5.54	0.71
	Curriculum vitae	4.05	1.68	3.93, 4.18	0.84
	Interview	5.17	1.32	5.07, 5.27	0.78
	Motivation letter	4.62	1.48	4.50, 4.73	0.72
	Personality questionnaire	3.87	1.59	3.75, 3.99	0.77
	Pre-university grade-point average	3.70	1.59	3.58, 3.82	0.81
	Situational judgement test	4.39	1.42	4.28, 4.50	0.77
	Skills test	5.27	1.20	5.18, 5.36	0.78
	Unweighted lottery	2.50	1.74	2.37, 2.63	0.75
	Weighted lottery	2.99	1.65	2.87, 3.12	0.80
Applicant differentiation	Cognitive capacity test	4.61	1.49	4.50, 4.73	0.57
	Curriculum-sampling test	4.59	1.45	4.48, 4.70	0.46
	Curriculum vitae	4.41	1.70	4.28, 4.54	0.72
	Interview	5.23	1.37	5.13, 5.33	0.68
	Motivation letter	4.35	1.64	4.22, 4.47	0.63
	Personality questionnaire	4.36	1.57	4.24, 4.48	0.64
	Pre-university grade-point average	3.25	1.55	3.13, 3.36	0.62
	Situational judgement test	4.24	1.54	4.13, 4.36	0.60
	Skills test	5.25	1.21	5.16, 5.34	0.57
Study relatedness	Cognitive capacity test	3.72	1.38	3.62, 3.82	0.56
	Curriculum-sampling test	4.38	1.32	4.28, 4.48	0.45
	Curriculum vitae	3.11	1.48	3.00, 3.22	0.66
	Interview	3.91	1.38	3.81, 4.01	0.57
	Motivation letter	3.32	1.43	3.21, 3.42	0.57
	Personality questionnaire	3.24	1.42	3.13, 3.34	0.63
	Pre-university grade-point average	2.93	1.47	2.82, 3.04	0.55
	Situational judgement test	3.57	1.46	3.46, 3.68	0.60
	Skills test	4.43	1.38	4.33, 4.53	0.50
Chance to perform	Cognitive capacity test	4.20	1.54	4.08, 4.31	0.59
	Curriculum-sampling test	4.58	1.35	4.48, 4.68	0.48
	Curriculum vitae	4.24	1.79	4.11, 4.38	0.65
	Interview	4.77	1.51	4.66, 4.89	0.61
	Motivation letter	4.23	1.57	4.11, 4.34	0.47
	Personality questionnaire	3.80	1.65	3.68, 3.92	0.50
	Pre-university grade-point average	3.09	1.61	2.97, 3.21	0.59
	Situational judgement test	3.66	1.57	3.55, 3.77	0.57
	Skills test	4.90	1.35	4.80, 5.00	0.50
Ease of cheating	Cognitive capacity test	2.17	1.37	2.07, 2.27	−0.11
	Curriculum-sampling test	2.29	1.48	2.17, 2.40	−0.28
	Curriculum vitae	4.42	1.85	4.28, 4.56	−0.24
	Interview	4.62	1.62	4.50, 4.74	−0.29
	Motivation letter	5.72	1.76	5.62, 5.82	−0.24
	Personality questionnaire	6.06	1.30	5.96, 6.16	−0.26
	Pre-university grade-point average	2.15	1.51	2.04, 2.26	−0.09
	Situational judgement test	4.72	1.92	4.57, 4.86	−0.25
	Skills test	2.98	1.50	2.83, 3.06	−0.21

Note: M = mean, SD = standard deviation, 95% CI = 95% confidence interval, R = correlation with general favourability,
[a]All correlations were statistically significant (p < 0.05).

TABLE 4 Overview of justice dimensions, codes and sample quotations of the qualitative content analyses

Justice dimension	Code	Sample quotation
Face validity	Combination of methods	This is a good method for selection, only if it is combined with a motivation letter on why you want to do the medical program and why you deserve to be admitted. Then you will also immediately learn something about the character and qualities of the person. (Curriculum-sampling test)
	Added value	I chose to enroll in a study program that did not ask about my grades. In my opinion, a pre-university diploma is already a good indication of a person's cognitive abilities. (Pre-university grade-point average)
Applicant differentiation	Distinctiveness	By assessing someone's knowledge, skills or opinion, you get a good idea of this person. Certain skills will emerge and each applicant can be distinguished. (Skills test)
		I find the use of a Situational Judgement Test as a selection method very bad. The situation as in the example item can be answered correctly by an average pre-university student in such a way that the 'desired' answer is given. If everyone gives the 'desired' answer, then no one is 'special' anymore, so you have to select again in a different way. (SJT)
	Origin of the score	What do grades say? Someone can put in a lot of effort and get a 7, and someone can put in a little effort and get a 7. (Pre-university grade-point average)
	Skills on which distinction is based	A motivation letter will roughly say something about the person, but I think you mainly see a great distinction based on who can and who cannot write well and not based on whether you are suitable for the medical program. A motivation letter would be perfect for journalism, but I do not think it should be decisive for selection to medicine. (Motivation letter)
	Format	It is often the case that a maximum number of characters is fixed for the motivation letter. This is actually the limiting factor to motivation letters. Some people have much more extensive motivations that are not apparent due to this limit. You quickly know how great a person's motivation is based on the size of the motivation letter itself. (Motivation letter)
Study relatedness	Relevance	A motivation letter enables the prospective student to describe the reason for his interest in medicine. The prospective student can show his social skills in writing. He must logically convey a reasoning in writing so everyone can understand it. Understanding of language and mindset become clear to the selection committee. I think this is essential for good medical students. (Motivation letter)
		A personality questionnaire does indeed say something about, for example, your communication skills or work attitude. However, it says nothing about your intelligence, learning capacity or quality of work. (Personality questionnaire)
		If you are active outside of school, are you also a better student? I do not think that has much to do with it. (Curriculum vitae)
	Specificity	Do not know if this is the best way. Cognitive ability test is of course not only suitable for pharmacy, this could also be taken as part of a psychology program. I would relate the selection more to pharmacy itself, although I think it is good that value is attached to reading comprehension and reasoning. (Cognitive capacity test)
	Learnability	These skills will also be taught in medical school. Now students are being tested on something that they actually have yet to learn. A bit unfair for those who do not have these skills yet. (Skills test)
	Preview of the programme	It's a good indication, since you have to study for a test on medicine, and so you get a little taste of the program. (Curriculum-sampling test)
	Job relatedness	It just shows how well one can learn and master the material. So, it can provide insight into how well someone studies and how good someone will be as a student. But it does not show how good a person is a s a doctor. (Curriculum-sampling test)
Chance to perform	Snapshot assessment	It is not fair, because if you have a bad day once and that is exactly on the testing day, you immediately lost your chance. It is fair to take several tests and look at their average. (Cognitive capacity test)
	Ability to show who you are	Really motivated candidates with their heart in the right place can prove themselves in an interview. (Interview)
	Effort expectancy	The fate of this test is really in your own hands. If you prepare well, you will make it, and if you do not prepare well, you will not make it. (Curriculum-sampling test)
	Accessibility[a]	In itself, it would be a suitable method, but people with, for example, doctors in the family or schools with many extra activities probably have more options because they have more connections. This does not mean that they are not actively trying to get in, but some people just do not have these options. (Curriculum vitae)
		Statistically, the number of selected men compared to the number of selected women will be much smaller. Men do not start to take school seriously until later in life (proven with research). (Pre-university grade-point average)
	Timing	It is better than lottery based on cognitive skills from year 5, because some people are late bloomers or have only recently decided that they want to follow a numerus fixus program, and have only just started

TABLE 4 (Continued)

Justice dimension	Code	Sample quotation
		working hard for it. Selecting based on recent findings seems more efficient and fairer to me. (Cognitive capacity test)
	Personal circumstances	Events may have occurred, as a result of which pre-university education did not go well (family problems, illness, etc.) which could affect year 5 grades. (Pre-university grade-point average)
	Preparation time	The available preparation time of candidates often differs too much. Some candidates are called in sick from school by their parents to learn, some have time to learn at school in certain classes, and others only have time to learn outside of school. (Curriculum-sampling test)
	Information	We were limitedly informed about the importance of good grades for the study, so many students did not try their best in year 4 and year 5. Many assumed that there was no selection but lottery. (Pre-university grade-point average)
Ease of cheating	Social desirability	I do not find a personality questionnaire useful as a selection method. If the applicant is well aware of the qualities needed for a doctor/medical student, he/she can complete the questionnaire in such a way that it looks like they are the perfect candidate. The objectivity of this selection method is lacking and, in my view, that is one of the most important aspects of a fair selection. (Personality questionnaire)
	Plagiarism and fraud	Lying on a CV is easy, but that is also fraud. If the selection committee also checks certain sources when checking the CV by calling or checking websites etc., they will automatically find out whether something is correct or not. So, in that respect, the CV is a fair method. (CV) From my own experience I often see that people have someone else write the motivation letter. With this method, it is difficult to check whether the words are the applicant's own words. (Motivation letter) Even if proof is requested, there is always someone willing to make proof of something that never happened. (Curriculum vitae)
Consistency of administration	Objectiveness	There is usually only one correct answer on such tests, so there is no room for subjectivity on part of the assessors. (Curriculum-sampling test)
	Consistency	Schools differ greatly in their assessment methods. As a result, it is relatively easier to get high grades in some schools than in other schools. (Pre-university grade-point average)
Reactions during hiring	Well-being	In my opinion, it is a strange idea that young people who, so to speak, should still play outside and go to football, should be concerned with, for example, gaining administrative experience. A candidate would not be ambitious enough if he or she has not attended a health care institution in addition to his or her high school period. What? Let children be children. They're busy enough already. (Curriculum vitae) Because some people have been dreaming of becoming doctors for years, there can sometimes be too much pressure when taking this test. As a result, some students may score poorly even though they are very well suited for this study program. (Curriculum-sampling test)
Distributive justice	Diversity	It ensures diversity within the study program. (Unweighted lottery)

[a]The code accessibility was further divided into the subcodes: gender, age, migration background, socio-economic status, social network in the medical field and disability.

to the open-ended questions. A first justice dimension was consistency of administration, which refers to the extent to which decision procedures are consistent and without bias across people and over time.[22] CVs, motivation letters, interviews and skills tests were considered as more subjective methods, giving a lot of space to the personal opinion of assessors and consequently creating potential differences in ratings between assessors. Applicants also commented on the consistency of administration for pu-GPA. Currently, pu-GPA is based on school-specific examinations. Applicants mentioned that there is a great difference between schools in assessment methods, difficulty level and way of rating, making pu-GPA from different schools incomparable.

Second, while the questionnaire only focused on study relatedness, applicants tended to make a distinction between study relatedness and job relatedness. For instance, while applicants had generally positive comments about curriculum sampling with respect to study

relatedness, they mentioned that other skills are relevant to become a successful health professional, such as social and communicative skills. They stated that skills tests, interviews, SJTs and personality questionnaires can assess important attributes for the future profession, while this was not so much the case for pu-GPA, cognitive tests, CVs and curriculum-sampling tests.

Some comments were not related to procedural justice, but rather to distributive justice, relating to the fairness of the outcomes of the selection procedure.[15] Applicants mentioned that an advantage of unweighted lottery is that the equal chances to be admitted would lead to a more diverse student population. The application of other methods could result in a decrease of student diversity, because of the aforementioned inequal chances to perform. An additional unfavourable outcome was mentioned for SJTs: accepted students would have similar opinions and there would be less diversity in perspectives in the student cohorts.

Finally, numerous remarks were made about the 'outcomes' part of organisational justice theory, which relates to the attitudes and behaviours of individuals that are thought to be a result of perceptions of selection.[15] One particular topic of concern was that selection methods can cause a lot of stress and pressure amongst applicants. Applicants mentioned that selection methods that require a lot of preparation, such as building a CV and studying for a curriculum-sampling test, can be stressful as they have to combine this with school and other responsibilities. Another source of stress that was mentioned was the pressure that applicants experience during testing days due to the high stakes involved. Applicants also recognised stress as a drawback of combining multiple selection methods.

3.4 | Differences in perceptions for subgroups of applicants

The results of the linear mixed model for subgroup differences in general favourability ratings for each selection method are depicted in Table 5, and the descriptive statistics for each subgroup are provided in supplemental Tables S3–S6. Compared with women, men perceived unweighted lottery as significantly less favourable ($B = -0.55$, 95% CI $[-0.84, -0.26]$), indicating that on average, men rated unweighted lottery 0.55 units lower on a scale from 1–7. Personality questionnaires were also perceived as significantly less favourable by men ($B = -0.23, [-0.47, -0.00]$). No gender differences were found for other selection methods. Interestingly, in the answers to open-ended questions, multiple remarks were given about CV and pu-GPA favouring women, describing that men generally start later with their future study orientation. This concern was not reflected in the quantitative findings.

Applicants with a non-Western migration background provided significantly higher favourability ratings for CVs ($B = 0.50 [0.22, 0.77]$) and curriculum-sampling tests ($B = 0.25 [0.06, 0.44]$), while skills tests and interviews were perceived significantly less favourable (respectively, $B = -0.29 [-0.50, -0.09]$, $B = -0.42 [-0.64, -0.20]$). The lower scores on interviews and skills tests for applicants with a non-Western migration background may be explained by remarks about the subjective nature of these methods, making them more susceptible to bias. Perceptions of applicants with a Western migration background did not differ from those without a migration background.

Compared with applicants applying during their final year of pre-university education, applicants who were already studying at university-level rated interviews and unweighted lottery significantly more positive (respectively, $B = 0.33 [0.07, 0.58]$, $B = 0.54 [0.17, 0.92]$), while their favourability ratings were significantly lower for pu-GPA ($B = -0.83 [-1.14, -0.52]$), cognitive tests ($B = -0.31 [-0.59, -0.04]$) and weighted lottery ($B = -0.41 [-0.74, -0.08]$). Applicants from alternative forms of prior education rated personality questionnaires and interviews significantly more favourable compared with applicants from pre-university education (respectively, $B = 0.49 [0.09, 0.89]$, $B = 0.46 [0.12, 0.80]$). Only with respect to pu-GPA, applicants commented on the role of prior education, indicating that pu-GPAs

for applicants not applying during their final year of pre-university education are outdated and do not accurately represent applicants' current skills.

No differences in applicant perceptions were found based on first-generation university status.

4 | DISCUSSION

The aim of the present study was to gain a deeper understanding of how applicants perceive different selection methods within the context of undergraduate HPE. Our findings indicate that applicants have the strongest preference for curriculum-sampling tests and skills tests, while they consider weighted and unweighted lotteries the least favourable. Furthermore, applicant perceptions of selection methods are overall similar across subgroups of applicants based on gender, ethnicity, SES and prior education. Finally, the results show that applicants do not think one single method is sufficient, but instead prefer a combination of selection methods.

A first key finding is that applicants have a preference for 'broadened' selection criteria and selection methods on which they feel to be more 'in control'. The three highest rating selection methods – curriculum-sampling tests, skills tests and interviews—all reflect broadened criteria, which aim to assess qualities that go beyond the cognitive abilities that are included in traditional methods such as pu-GPA.[23] Applicants in the current study were already strongly preselected based on cognitive characteristics due to the stringent admission requirements; thus, it is not surprising that they prefer to distinguish themselves in other areas. In fact, applicants questioned the added value of pu-GPA and cognitive capacity tests over the admission requirements. Additionally, according to organisational justice theory,[15] as well as previous research in medical school selection,[24] applicants have the desire to 'express themselves' during a selection procedure, and it can be argued that this is better possible with methods that reflect broadened criteria. The low ratings of lotteries—leaving it to chance—and pu-GPA—obtained prior to selection—indicate that applicants prefer to be 'in control'. As previously argued,[1,2] this can also be explained through organisational justice theory, because applicants perceive methods as fairer when they have an opportunity to influence the decision process.[15] This finding can furthermore be related to a broader societal appreciation of meritocratic values, implying that success and failure can be attributed to one's own efforts and talent.[25] Consequently, applicants for selection into HPE can find rejections that feel beyond their control difficult to accept.[26]

Second, our findings suggest that applicants' background does not play a substantial role in their perceptions of different selection methods. Although we did find some statistically significant differences in perceptions between applicant subgroups based on gender, migration background and prior education, the practical meaning of these differences is negligible, and their perceptions were overall very similar. For gender, this lack of difference in perceptions is in accordance with findings of a previous study conducted with psychology

TABLE 5 Linear mixed model results for subgroup differences in general favourability ratings for each selection method

	Parameter	B	Standard error	95% CI	t
Main effects	Intercept (unweighted lottery)	3.04***	0.10	2.84, 3.24	29.70
	Cognitive capacity test	1.67***	0.13	1.42, 1.92	13.23
	Curriculum-sampling test	2.24***	0.12	2.00, 2.48	18.22
	Curriculum vitae	0.73***	0.14	0.46, 1.01	5.30
	Interview	2.07***	0.13	1.82, 2.32	16.41
	Motivation letter	1.52***	0.13	1.27, 1.78	11.67
	Personality questionnaire	0.87***	0.13	0.61, 1.12	6.71
	Pre-university grade-point average	0.51***	0.14	0.24, 0.78	3.73
	Situational judgement test	1.32***	0.13	1.07, 1.57	10.47
	Skills test	2.17***	0.12	1.93, 2.40	17.99
	Weighted lottery	0.12	0.13	−0.14, 0.37	0.91
Selection method × gender	Cognitive capacity test * man	0.21	0.11	−0.01, 0.42	1.88
	Curriculum-sampling test * man	0.01	0.09	−0.17, 0.18	0.07
	Curriculum vitae * man	−0.07	0.13	−0.32, 0.19	−0.50
	Interview * man	−0.04	0.10	−0.24, 0.16	−0.43
	Motivation letter * man	−0.21	0.11	−0.43, 0.01	−1.84
	Personality questionnaire * man	−0.23*	0.12	−0.47, −0.00	−1.99
	Pre-university grade-point average * man	0.17	0.12	−0.08, 0.41	1.34
	Situational judgement test * man	−0.13	0.11	−0.34, 0.09	−1.17
	Skills test * man	−0.10	0.10	−0.29, 0.10	−0.98
	Unweighted lottery * man	−0.55***	0.15	−0.84, −0.26	−3.71
	Weighted lottery * man	0.03	0.13	−0.23, 0.29	0.22
Selection method × migration background	Cognitive capacity test * Western	−0.10	0.18	−0.45, 0.25	−0.55
	Cognitive capacity test * non-Western	−0.20	0.12	−0.43, 0.03	−1.68
	Curriculum-sampling test * Western	−0.03	0.15	−0.31, 0.26	−0.18
	Curriculum-sampling test * non-Western	0.25*	0.10	0.06, 0.44	2.53
	Curriculum vitae * Western	0.14	0.21	−0.28, 0.55	0.66
	Curriculum vitae * non-Western	0.50***	0.14	0.22, 0.77	3.54
	Interview * Western	−0.16	0.16	−0.48, 0.17	−0.95
	Interview * non-Western	−0.42***	0.11	−0.64, −0.20	−3.80
	Motivation letter * Western	−0.23	0.18	−0.59, 0.13	−1.24
	Motivation letter * non-Western	0.00	0.12	−0.24, 0.24	0.00
	Personality questionnaire * Western	−0.12	0.19	−0.50, 0.26	−0.61
	Personality questionnaire * non-Western	−0.20	0.13	−0.45, 0.05	−1.57
	Pre-university grade-point average * Western	0.11	0.20	−0.29, 0.50	0.54
	Pre-university grade-point average * non-Western	0.16	0.13	−0.10, 0.43	1.21
	Situational judgement test * Western	−0.09	0.17	−0.43, 0.25	−0.50
	Situational judgement test * non-Western	0.10	0.12	−0.13, 0.33	0.84
	Skills test * Western	−0.01	0.16	−0.32, 0.30	−0.08
	Skills test * non-Western	−0.29**	0.11	−0.50, −0.09	−2.79
	Unweighted lottery * Western	−0.18	0.24	−0.65, 0.29	−0.75
	Unweighted lottery * non-Western	0.00	0.16	−0.31, 0.32	0.02
	Weighted lottery * Western	−0.34	0.21	−0.76, 0.08	−1.58
	Weighted lottery * non-Western	0.01	0.14	−0.27, 0.29	0.05
Selection method × prior education	Cognitive capacity test * university	−0.31*	0.14	−0.59, −0.04	−2.23
	Cognitive capacity test * other	0.10	0.19	−0.27, 0.47	0.54

(Continues)

TABLE 5 (Continued)

	Parameter	B	Standard error	95% CI	t
	Curriculum-sampling test * university	0.16	0.12	−0.06, 0.39	1.41
	Curriculum-sampling test * other	0.03	0.15	−0.27, 0.33	0.17
	Curriculum vitae * university	−0.05	0.17	−0.37, 0.28	−0.28
	Curriculum vitae * other	0.24	0.22	−0.19, 0.67	1.09
	Interview * university	0.33[*]	0.13	0.07, 0.58	2.51
	Interview * other	0.46[**]	0.17	0.12, 0.80	2.65
	Motivation letter * university	0.10	0.15	−0.19, 0.38	0.68
	Motivation letter * other	0.38	0.19	0.00, 0.75	1.96
	Personality questionnaire * university	−0.07	0.15	−0.36, 0.23	−0.46
	Personality questionnaire * other	0.49[*]	0.20	0.09, 0.89	2.41
	Pre-university grade-point average * university	−0.83[***]	0.16	−1.14, −0.52	−5.24
	Pre-university grade-point average * other	0.20	0.21	−0.22, 0.62	0.94
	Situational judgement test * university	0.05	0.14	−0.22, 0.33	0.39
	Situational judgement test * other	0.14	0.18	−0.21, 0.50	0.79
	Skills test * university	0.22	0.13	−0.03, 0.46	1.74
	Skills test * other	0.04	0.17	−0.29, 0.37	0.24
	Unweighted lottery * university	0.54[*]	0.19	0.17, 0.92	2.86
	Unweighted lottery * other	0.40	0.25	−0.10, 0.89	1.56
	Weighted lottery * university	−0.41[*]	0.17	−0.74, −0.08	−2.42
	Weighted lottery * other	0.30	0.23	−0.15, 0.74	1.32
Selection method × parental education	Cognitive capacity test * 1st generation	−0.07	0.12	−0.30, 0.16	−0.62
	Curriculum-sampling test * 1st generation	−0.14	0.10	−0.33, 0.05	−1.46
	Curriculum vitae * 1st generation	0.15	0.14	−0.12, 0.45	1.07
	Interview * 1st generation	−0.08	0.11	−0.29, 0.14	−0.68
	Motivation letter * 1st generation	0.03	0.12	−0.21, 0.27	0.21
	Personality questionnaire * 1st generation	0.03	0.13	−0.22, 0.28	0.21
	Pre-university grade-point average * 1st generation	−0.09	0.13	−0.36, 0.17	−0.69
	Situational judgement test * 1st generation	−0.01	0.12	−0.24, 0.22	−0.06
	Skills test * 1st generation	0.01	0.11	−0.19, 0.22	0.13
	Unweighted lottery * 1st generation	−0.04	0.16	−0.36, 0.27	−0.27
	Weighted lottery * 1st generation	−0.22	0.14	−0.50, 0.06	−1.55

Note: B refers to the unstandardised regression coefficient together with the standard error and the 95% confidence interval (95% CI). Reference categories: unweighted lottery for main effects, female gender, standard pre-university education, no migration background and no first-generation university applicant for interaction effects. Random effects: programme, participant ID.

[*]$p < 0.05$.
[**]$p < 0.01$.
[***]$p < 0.001$.

applicants.[2] However, our findings with respect to ethnicity and SES were unexpected, as previous research suggested that students with an ethic minority or lower socio-economic background have more negative perceptions towards selection.[1,3,11] Potentially, applicants perceive the formal characteristics of selection (i.e. selection methods) similar regardless of their background, while applicants with sociodemographic minority backgrounds may have more negative perceptions regarding other components of procedural justice, such as interpersonal treatment. Indeed, research has indicated that ethnic minority students have more perceptions of unfair treatment when they are in medical school.[27,28] An alternative reason for the finding that applicants with minority backgrounds did not have more negative perceptions towards certain selection methods is that previous studies mainly included eligible pre-university students who have not yet applied and could have based their perceptions on inaccurate or incomplete information.[3,11] Contrarily, participants in the present study had experience with selection and probably had a better understanding of it. Thus, students with minority backgrounds possibly

experience less access to accurate information about the selection procedure mainly in the period *prior* to application.[29] Nevertheless, corresponding with previous studies, the qualitative findings did provide indications that applicants believe selection methods can hinder equitable admission to medical education, for instance, because applicants with a minority background have less access to coaching, tutoring and a social network in the medical field.[3,11,29]

A third key finding is that according to applicants, there is no ultimate solution; all selection methods are accompanied by certain advantages and disadvantages. This is well reflected by a number of conflicting preferences that we identified in the qualitative results. The most distinctive dilemmas we found were (1) measuring what applicants can offer at the present versus countering negative consequences of snapshot assessments, (2) measuring existing knowledge and skills versus providing space to grow and develop skills during the program, (3) selecting the best students versus selecting the best future professionals and (4) combining multiple selection methods for comprehensiveness versus preventing too much stress and pressure. The first dilemma evidently reflects the justice dimension of chance to perform and stresses the aforementioned desire of applicants to feel in control of the selection process. The second and third dilemmas relate to study and job relatedness and are also topic of debate within the academic field.[23] The final dilemma relates more to another part of organisational justice theory, namely, the relationship between applicants' perceptions and individual and organisational outcomes.[15] According to the theory, perceptions of fairness can affect variables such as motivation, self-esteem and self-efficacy (referred to as 'reactions during hiring').[15] Feelings of stress and pressure were brought up as a topic of concern in the present study, which applicants mainly related to the amount of preparation and the high stakes involved. A review also identified considerable prevalence of psychological stress amongst students in HPE.[30] Nevertheless, thus far, studies have paid little to no attention to applicant well-being with respect to selection and the particular concept of stress is not included in organisational justice theory. Results of the present study indicate that it may be relevant to add this to the theory as an additional subcategory of 'reactions during hiring', although further research would be required to validate this.

A strength of the present study is that, to our knowledge, it is the first to compare perceptions of different sociodemographic subgroups and seek for the underlying argumentation behind applicant perceptions for a great range of selection methods. Additionally, we collected data from multiple programmes. Consequently, applicants also provided their opinions for methods that they had not been exposed to. This way, our results were less influenced by self-selection, given that the selection procedure can play a role in applicants' programme of choice.[16] However, this strength was also accompanied by the limitation that the respondents did have experience with some of the methods in the questionnaire, while for other methods, their perceptions were based solely on a short description of the selection method. Some of the methods were not employed by any of the programmes. We mitigated this as much as possible by controlling for the programme to which applicants had applied. Furthermore, the survey

that was administered did not capture all factors that could influence applicant perceptions, such as comprehensiveness, defensibility and the role of the process.[31] Another limitation of the present study was that, although the sample size was relatively large, only 21% of invited applicants responded to the survey. Consequently, their perceptions may not be fully representative for the complete applicant pool. Nevertheless, our sample was comparable with applicant pools in previous studies in terms of demographic characteristics.[12,17] A final limitation is that the present study could not directly compare the argumentation of different subgroups of applicants due to the voluntary nature of the open-ended questions and the unequal distribution of subgroups. Consequently, certain subgroup differences in perceptions that could not be captured by the quantitative data may have remained undiscovered.

The current study focused on one aspect of procedural justice (i.e. formal characteristics), but future studies can also pay attention to other justice components, such as interpersonal treatment.[15] Additionally, more fundamental questions regarding applicant perceptions have not been investigated yet, such as: What mechanisms play a role in the development of these perceptions? What are the underlying values of applicants on which their perceptions are based? Furthermore, in the present study, applicants preferred a combination of different methods, so it could be valuable to examine how they perceive selection procedures with different combinations of methods. It would be particularly interesting to evaluate applicant perceptions regarding a holistic approach, as this approach pays attention to both comprehensiveness and WA.[32] Furthermore, results of the present study suggest that applicants have a stronger preference for the methods employed by the programmes to which they have applied and for which they have prepared. Future research could examine whether applicant perceptions of the hidden curriculum play a role in this preference. Another interesting direction for future research relates to the question of whether the relationship between perceptions and performance during selection differs based on applicants' background. Previous studies have indicated that applicant perceptions and performance are positively correlated,[2,5,6] and that certain demographic subgroups perform less well during selection,[12,13,17,18] while we found that perceptions were similar for applicants regardless of their background. Additionally, it can be relevant to investigate subgroup perceptions for eligible students who have not applied yet. Finally, researchers can consider including items on consistency in administration and job relatedness, as well as applicant well-being, because these topics were considered relevant by applicants.

From a practical viewpoint, our findings provide insights that can help selection committees design their procedures, by considering selection methods that are preferred by applicants. However, there can be frictions between applicant perceptions of selection methods (an element of political validity) and other indicators of validity.[1,2] For instance, while applicants hold negative perceptions towards the use of pu-GPA as a selection method, pu-GPA is strongly predictive of future academic performance.[33] Likewise, the findings of the present study suggest that the inclusion of curriculum-sampling tests can attract more applicants with a non-Western background, but we

found in a submitted study that this subgroup of applicants performs less well on such tests.[34] Thus, when deciding which selection methods to include, applicant perceptions can be taken into account, while also considering other aspects of validity in a broad sense.[35] Nevertheless, our findings can help identify which selection methods are less attractive and require better explanation.[2] Clear and transparent communication about selection methods may improve applicant perceptions, especially in the case of broadened criteria that are perceived as more 'subjective' and for which perceptions were sometimes based on inaccurate information. For instance, applicants believed that a social network in the medical field is necessary for building a CV, while the programmes in our study also value other relevant experience outside of the medical field. Although understanding applicant perceptions cannot provide clear-cut solutions about which selection methods programmes should include, our qualitative findings provide valuable insights into how programmes can adjust the implementation of their selection methods to take applicants' needs into account to improve motivation and performance or prevent withdrawal. For example, because applicants were concerned about the impact of selection on their well-being, programmes could limit the volume of preparatory materials for curriculum-sampling tests.

In conclusion, applicants participating in selection for undergraduate HPE prefer selection methods for which they perceive to be in control and which assess other qualities than cognitive ability. Additionally, the present study indicated that applicant perceptions of individual selection methods are generally similar between applicants with majority and minority backgrounds. Due to contradictory needs of applicants and frictions between applicant perceptions and other indicators of validity, we call for a clearer and more transparent communication of selection methods and modifications *within* methods.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

None.

ETHICS STATEMENT

The present study was carried out in accordance with the Declaration of Helsinki and was deemed exempt from full review after evaluation by the Medical Ethics Committee of Erasmus University Medical Center, Rotterdam.

AUTHOR CONTRIBUTIONS

SFW, KMSJ and AMW substantially contributed to the conception and design of the study. KMSJ, RAVG, MG, JHR and AW were responsible for data collection. SFW analysed the data and wrote the first draft of the article. All authors interpreted the data, revised it critically for intellectual content, approved the final manuscript for publication and have agreed to be accountable for all aspects of the work in ensuring that questions related to its accuracy or integrity are appropriately investigated and resolved.

ORCID

Suzanne Fikrat-Wevers https://orcid.org/0000-0001-8022-3746
Karen Stegers-Jager https://orcid.org/0000-0003-2947-6099
Marleen Groenier https://orcid.org/0000-0001-7298-8431
Andries Koster https://orcid.org/0000-0001-7772-4654
Anouk Wouters https://orcid.org/0000-0002-6806-7573

REFERENCES

1. Kelly ME, Patterson F, O'Flynn S, Mulligan J, Murphy AW. A systematic review of stakeholder views of selection methods for medical schools admission. *BMC Med Educ*. 2018;18(1):1-26. doi:10.1186/S12909-018-1235-X/TABLES/1
2. Niessen ASM, Meijer RR, Tendeiro JN. Applying organizational justice theory to admission into higher education: admission from a student perspective. *Int J Sel Assess*. 2017;25(1):72-84. doi:10.1111/ijsa.12161
3. Greenhalgh T, Seyan K, Boynton P. "Not a university type": focus group study of social class, ethnic, and sex differences in school pupils' perceptions about medical school. *BMJ*. 2004;328(7455):1541. doi:10.1136/BMJ.328.7455.1541
4. Cleland J, Dowell J, Mclachlan J, Nicholson S, Patterson F. *Identifying Best Practice in the Selection of Medical Students (Literature Review and Interview Survey)*. General Medical Council; 2012.
5. Chan D, Schmitt N. Video-based versus paper-and-pencil method of assessment in situational judgment tests: subgroup differences in test performance and face validity perceptions. *J Appl Psychol*. 1997;82(1):143-159. doi:10.1037/0021-9010.82.1.143
6. Truxillo DM, Steiner DD, Gilliland SW. The importance of organizational justice in personnel selection: defining when selection fairness really matters. *Int J Sel Assess*. 2004;12(1-2):39-53. doi:10.1111/J.0965-075X.2004.00262.X
7. Thorsteinson TJ, Ryan AM. The effect of selection ratio on perceptions of the fairness of a selection test battery. *Int J Sel Assess*. 1997;5(3):159-168. doi:10.1111/1468-2389.00056
8. Yang L, Chang I, Ritz S, Grierson L. Research experiences for Canadian aspiring physicians: a descriptive analysis of medical school admission policies concerning research involvement in Canada. *BMC Med Educ*. 2022;22(1):1-10. doi:10.1186/S12909-022-03207-Y/TABLES/1
9. Lin KY, Anspach RR, Crawford B, Parnami S, Fuhrel-Forbis A, De Vries RG. What must I do to succeed?: narratives from the US premedical experience. *Soc Sci Med*. 2014;119:98-105. doi:10.1016/J.SOCSCIMED.2014.08.017
10. White J, Brownell K, Lemay JF, Lockyer JM. "What do they want me to say?" the hidden curriculum at work in the medical school selection process: a qualitative study. *BMC Med Educ*. 2012;12(1):17. doi:10.1186/1472-6920-12-17
11. Wouters A, Croiset G, Isik U, Kusurkar RA. Motivation of Dutch high school students from various backgrounds for applying to study medicine: a qualitative study. *BMJ Open*. 2017;7(5):e014779. doi:10.1136/bmjopen-2016-014779
12. Mulder L, Wouters A, Twisk JWR, et al. Selection for health professions education leads to increased inequality of opportunity and decreased student diversity in The Netherlands, but lottery is no solution: a retrospective multi-cohort study. *Med Teach*. 2022;44(7):790-799. doi:10.1080/0142159X.2022.2041189
13. Mathers J, Sitch A, Parry J. Population-based longitudinal analyses of offer likelihood in UK medical schools: 1996-2012. *Med Educ*. 2016;50(6):612-623. doi:10.1111/medu.12981
14. De Leng WE, Stegers-Jager KM, Born MP, Themmen APN. Influence of response instructions and response format on applicant perceptions of a situational judgement test for medical school selection. *BMC Med Educ*. 2018;18(1):282. doi:10.1186/s12909-018-1390-0

15. Gilliland SW. The perceived fairness of selection systems: an organizational justice perspective. *Acad Manag Rev*. 1993;18(4):694-734. doi:10.5465/amr.1993.9402210155

16. Wouters A, Croiset G, Schripsema NR, et al. Students' approaches to medical school choice: relationship with students' characteristics and motivation. *Int J Med Educ*. 2017;8:217-226. doi:10.5116/ijme.5921.5090

17. Stegers-Jager KM, Steyerberg EW, Lucieer SM, Themmen APN. Ethnic and social disparities in performance on medical school selection criteria. *Med Educ*. 2015;49(1):124-133. doi:10.1111/medu.12536

18. Van Den Broek A, Mulder J, De Korte K, Bendig-Jacobs J, Van Essen M. *Selectie Bij Opleidingen Met Een Numerus Fixus & de Toegankelijkheid van Het Hoger Onderwijs*. ResearchNed; 2018.

19. Niessen ASM, Meijer RR. Selection of medical students on the basis of nonacademic skills: is it worth the trouble? *Clin Med*. 2016;16(4):339-342. doi:10.7861/clinmedicine.16-4-339

20. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res*. 2005;15(9):1277-1288. doi:10.1177/1049732305276687

21. Zhang Y, Wildemuth BM. Qualitative analysis of content. In: Wildemuth BM, ed. *Applications of Social Research Methods to Questions in Information and Library Science*. Libraries Unlimited; 2009:308-319.

22. Bauer TN, Truxillo DM, Sanchez RJ, Craig JM, Ferrara P, Campion MA. Applicant reactions to selection: development of the selection procedural justice scale (SPJS). *Pers Psychol*. 2001;54(2):387-419. doi:10.1111/J.1744-6570.2001.TB00097.X

23. Stegers-Jager KM. Lessons learned from 15 years of non-grades-based selection for medical school. *Med Educ*. 2018;52(1):86-95. doi:10.1111/medu.13462

24. White JS, Lemay JF, Brownell K, Lockyer J. "A chance to show yourself"—how do applicants approach medical school admission essays? *Med Teach*. 2011;33(10):e541-e548. doi:10.3109/0142159X.2011.599890

25. Sandel MJ. *Tyranny of Merit: Can We Find the Common Good?* Picador; 2021.

26. Ten Cate O. Rationales for a lottery among the qualified to select medical trainees: decades of Dutch experience. *J Grad Med Educ*. 2021;13(5):612-615. doi:10.4300/JGME-D-21-00789.1

27. van Andel CEE, Born MP, van den Broek WW, Stegers-Jager KM. Student ethnicity predicts social learning experiences, self-regulatory focus and grades. *Med Educ*. 2022;56(2):211-219. doi:10.1111/MEDU.14666

28. Isik U, Wouters A, Verdonk P, Croiset G, Kusurkar RA. "As an ethnic minority, you just have to work twice as hard." Experiences and motivation of ethnic minority students in medical education. *Perspect Med Educ*. 2021;10(5):272-278. doi:10.1007/S40037-021-00679-4/TABLES/1

29. Martin AJ, Beska BJ, Wood G, et al. Widening interest, widening participation: factors influencing school students' aspirations to study medicine. *BMC Med Educ*. 2018;18:117. doi:10.1186/S12909-018-1221-3

30. Hope V, Henderson M. Medical student depression, anxiety and distress outside North America: a systematic review. *Med Educ*. 2014;48(10):963-979. doi:10.1111/MEDU.12512

31. Marrin ML, McIntosh KA, Keane D, Schmuck ML. Use of the paired-comparison technique to determine the most valued qualities of the McMaster medical programme admissions process. *Adv Heal Sci Educ*. 2004;9(2):129-135. doi:10.1023/B:AHSE.0000027439.18289.00

32. Conrad SS, Addams AN, Young GH. Holistic review in medical school admissions and selection: a strategic, mission-driven response to shifting societal needs. *Acad Med*. 2016;91(11):1472-1474. doi:10.1097/ACM.0000000000001403

33. Stegers-Jager KM, Themmen APN, Cohen-Schotanus J, Steyerberg EW. Predicting performance: relative importance of students' background and past performance. *Med Educ*. 2015;49(9):933-945. doi:10.1111/MEDU.12779

34. Fikrat-Wevers S, Stegers-Jager KM, Afonso PM, et al. Selection tools and student diversity in health professions education: a multi-site study. Submitted publication.

35. Schreurs S. Selection for Medical School: The Quest for Validity. Doctoral Thesis. Maastricht University; 2020. doi:10.26481/DIS.20200320SS

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.