The promise of potential: A study on the effectiveness of jury selection to a prestigious visual arts program

Kackovic, M.; Hartog, J.; van Ophem, H.; Wijnberg, N.

DOI
10.1111/kykl.12303

Publication date
2022

Document Version
Final published version

Published in
Kyklos

Received: 17 April 2020 | Revised: 10 December 2021 | Accepted: 4 February 2022

DOI: 10.1111/kykl.12303

ORIGINAL ARTICLE

KYKLOS WILEY

The promise of potential: A study on the effectiveness of jury selection to a prestigious visual arts program

Monika Kackovic[1] | Joop Hartog[2] | Hans van Ophem[3] | Nachoem Wijnberg[4]

[1]Amsterdam School of Business, University of Amsterdam, Amsterdam, the Netherlands

[2]University of Amsterdam, Amsterdam, the Netherlands

[3]Amsterdam School of Economics, University of Amsterdam and Tinbergen Institute, Amsterdam, the Netherlands

[4]Amsterdam School of Business, University of Amsterdam and University of Johannesburg, Johannesburg, South Africa

Correspondence
Monika Kackovic, Amsterdam School of Business, University of Amsterdam, Amsterdam, the Netherlands.
Email: m.kackovic@uva.nl

Abstract

We analyze 11 years of admission decisions to a highly selective post-graduate visual arts program. Our unique, longitudinal data includes detailed information about the selection procedures as well as the later artistic and economic achievements for both accepted and rejected applicants ($n = 8{,}557$). Regarding the predictive value of the selection procedures on later performance, our analyses show that the largest gain made is in the first, more cursory, pre-selection round. We find that weeding out less promising applicants is easier than identifying later top-performers, in terms of ex-ante expectations. The subsequent or final selection round is resource intensive and consists of several interviews with multiple jury members. Here we uncover an implicit structure of criterion variables that helps explain the ex-post admission decisions. Our simulations show that actual selection during the final round is slightly better than alternative decision rules (including a decision by lottery) in identifying applicants' potential. Nevertheless, measured by future performance, the gain from rigorous jury assessments relative to modest and cheaper selection methods is minimal. Our conclusions are maintained under several robustness checks.

1 | INTRODUCTION

Selection procedures focused on choosing the right candidate from a pool of applicants are common in many settings, ranging from hiring employees to choosing participants for academic or professional programs. Most selection decisions are made under conditions of uncertainty because information about the present characteristics of the applicants is imperfect and even more so about their potential in an uncertain future environment. Extant research has focused on formalized selection procedures using standardized test scores (Ehrenberg & Mavros, 1995; Grove & Wu, 2007), general mental ability tests (Ones et al., 2010; Schmitt, 2014) and measures rooted in formal structures that are non-binding or based on expert opinions (Glejser & Heyndels, 2001; Ginsburgh & Van Ours, 2003). We add to this literature by studying a specific kind of selection procedure, one which is unstructured and where the criteria for selection are not explicitly formulated, and the decision process of the jury is not formalized. The value of our research in analyzing this kind of selection is that it is a representation of a widely-used practice. We use real-world data that are rarely available to researchers - namely elaborate selection and career performance data of accepted and rejected applicants – to study the predictive value of such selection procedures.

The empirical setting of our study is the *Rijksakademie van Beeldende Kunsten* (RABK), an internationally renowned and prestigious post-graduate visual arts program located in the Netherlands. Their mission is to select the most talented visual artists and offer them a platform where research, experimentation, innovation, and critical discourse are central. It is important to note that this program does not focus on educating but rather provides facilities and networking opportunities for visual artists to expand their autonomous practice similar to, for instance, an academic post-doc position. To fulfill their goal, RABK invests heavily in talent scouting and development, starting with an annual rigorous multi-round selection process. During our 11-year observation period, 8,557 visual artists worldwide applied to the program and 3.5% were accepted. A unique feature of our data is that we are able to analyze the selection procedure and later market performance – measured, first and foremost, by artistic prestige rankings - of accepted and rejected applicants. Prestige rankings are a particularly useful performance metric for artists that are mostly active on the primary art market, where artworks are sold to the first buyer either directly by the artist or through an art gallery (Singer & Lynch, 1994). To address possible selectivity bias in our empirical analysis, we use semi-parametric switching regressions models based on Heckman (1974, 1976) and Newey (2009).

Our study contributes to the literature in three ways. First, we add to empirical studies on selective admission to highly prestigious programs (Ehrenberg & Mavros, 1995; Grove & Wu, 2007) based on expert opinions (Glejser & Heyndels, 2001; Ginsburgh & Van Ours, 2003) and informal interviews (Dana et al., 2013; Kausel et al., 2016) in a setting with large uncertainty about applicants' potential. Second, we uncover implicit selection criteria that have been used in assessing applicants during the selection rounds. We highlight the role of individual characteristics in a selection procedure to an internationally renowned art residency program that - if the applicant is accepted - has been shown to positively influence their career performance (Frey, 2019). Third, we extend the research on the labor supply of artists (e.g., Bille et al., 2017; Hartog & Kackovic, 2019) active in the contemporary art market, which is characterized by high rates of artistic innovation, oversupply, uncertain demand, and subjective quality evaluations (e.g., Caves, 2000; DiMaggio, 1987; Frey & Pommerehne, 1989; Velthuis, 2013). In addition, we present methodological insights concerning selection bias and simulate alternative decision rules.

This paper is organized as follows. Section 2 reviews the relevant literature on selection procedures for academic programs and employment. Section 3 provides an account of the selection procedures and data. Section 4 presents the primary analysis of the pre-selection round and the final selection round. Section 5 describes the latent structure of the criterion variables. Section 6 presents the estimates from the switching regressions. Section 7 discusses the main findings and Section 8 concludes.

2 | PREDICTIVE VALUE OF SELECTION PROCEDURES: A REVIEW

The literature in labor economics, occupational psychology, and human resource management point to the fact that forecasting future career success vis-à-vis selection procedures is difficult. Notwithstanding, we focus on two broad streams of research that addresses the predictive value of selection– namely, admission to and completion of an academic program or employment - on future career performance. The first is rooted in formalized structures, such as standardized test scores, coupled with structured interviews while the second focuses on predicting future performance based on general mental ability tests and unstructured interviews.

In labor economics, there is a stream of literature on predicting academic success of applicants to highly selective post-graduate academic programs. For instance, Ehrenberg and Mavros (1995) study the influence of Graduate Record Examinations (GRE) aptitude scores on completing a doctoral program in economics, English, mathematics and physics at a top-tier university. They found that, in general, students' ability, as measured by their GRE verbal and mathematics scores, is not associated with completing the program. One notable exception is economics, where higher verbal scores are associated with a higher likelihood of program completion. The authors note the weak relationship between GRE scores and applicants' "true ability." They suggest using additional benchmarks to measure quality, such as undergraduate school prestige, grade point average, endorsements, and admissions committee rankings to gain more valuable information about applicants' future potential.

Building upon the above mentioned research, Grove and Wu (2007) study the predictive value of objective and subjective applicant information on admissions to – and completion of – a top-tier economics doctoral program as well as individual future (publishing) performance after graduating the program. On the one hand, they found that verbal and mathematics GRE scores have a significant positive effect on the probability of achieving a doctorate degree, but only GRE mathematics scores predict future publication performance, i.e., a 50-point increase (from a mean value of 740) corresponds to a 7% increase in the likelihood of publishing. They also found that endorsements, i.e., letter of recommendation from prominent and prolific researchers, substantially increase the likelihood of acceptance as well as the probability of having at least one publication after finishing the program. Their estimations, however, show that graduating from an elite liberal arts or research program does not significantly influence doctoral program completion[1] or future publications. On the other hand, subjective ratings by the admissions committee members are also found to be predictors of doctoral completion and future publications. The authors conclude that when taken alone, a regression model containing objective information had better predictive power than a model with only subjective information. However, combing both kinds of applicant information (as suggested by Ehrenberg & Mavros, 1995) provided superior predictions.

Studies focusing exclusively on objective selection criteria on future employment performance have provided some evidence that general mental ability scores are predictors of applicants' potential (Ones et al., 2010; Schmitt, 2014), although the variance explained is not high (Cascio, 1999). Other studies focusing largely on reference letters from renowned and productive researchers find strong empirical support on the admissions decision (Castilla & Rissing, 2019). However, aside from the research conducted by Grove and Wu (2007) there is very little evidence that an endorsement significantly predicts academic excellence after admissions or future job market performance. In fact, Castilla and Rissing (2019) show that while highly recommended candidates who applied to a top MBA program at a prestigious university were more likely to pass the initial screening process, they received lower competency assessments during the admission interviews compared to those who were not recommended. Their results show that endorsed applicants did not perform academically more outstandingly or receive a higher salary after graduating than their non-endorsed peers. The authors argue that the advantage, which relies on applicants' social connections, is only beneficial during the initial or screening phase of application.

[1]A notable exception is Booth and Satchell (1995), who empirically demonstrate that graduating with honors from an elite undergraduate institution increases completion rates for females, although it has no effect on males.

Hence, compared to their non-endorsed counterparts, endorsed applicants are not "better performers" after being admitted to the program. While the results of these studies are not unequivocal, they do point to the usefulness - albeit in varying degrees - of objective and subjective applicant information on admissions decisions and future career performance.

Occupational psychology research and human resource management literature have shown that subjective selection committee rankings obtained through unstructured interviews predict future job performance imperfectly (Dana et al., 2013; McDaniel et al., 1994; Wiesner & Cronshaw, 1988; Wright et al., 1989). Nevertheless, this kind of selection process is widely practiced. Past research has shown that selection committees tend to overestimate the validity of unstructured interviews (Highhouse, 2008; Kausel et al., 2016; Lievens et al., 2005; Rynes et al., 2002; Terpstra, 1996). There appears to be a widely held perception that unstructured selection processes provide a window into capturing the unexplained variance of the applicant's future potential that is not attributable to other objectively measured competencies (Highhouse, 2008).

As this review of the literature indicates, selection for future success is not an easy task. No doubt, this undertaking is made even more difficult when assessing the potential of artists, whose output by definition is innovative and where subjective evaluations prevail (e.g., Caves, 2000). We focus on this new domain where unstructured selection procedures are relied upon to evaluate the hard-to-discern developmental potential and future performance of the applicants, particularly in light of the lack of availability of objective information, such as test scores or grades.

3 | THE SELECTION PROCEDURE OF THE PRESENT STUDY

The admission process for RABK has always been divided into two rounds, pre-selection and final selection. The pre-selection round leads up to an invitation for an interview, while during the final selection round admission decisions are made based upon the outcome of the interviews.[2] During pre-selection, jury members view all artworks in a digital format. They scan these and evaluate them based on technical skill and autonomous style. Moreover, the applicants' written description of the artworks and their expectations from the program are considered. Additionally, fit into the target group – i.e., age between 27 to 36 years old, at least 2-years independent art practice experience, and a letter of recommendation – is assessed. These specifications, however, are not rigidly applied as all applications are scrutinized for their unique qualifications.

In the final selection round, applicants are interviewed. During our observation period, there were two groups of jury members who conducted the interviews and a facilitator who took detailed notes. From 1994 to 1998, one group consisted of jurors specializing in two-dimensional artwork - e.g., painting, drawing, photography, graphic design - and the other in three-dimensional artwork - e.g., sculpture, installations - but film and art videos were also included. During this time frame, the jury of the core artistic discipline of the applicant made the final admittance decision. However, starting in 1999 up to the end of our observation period, the jury consisted of two multi-disciplinary groups, with four to five internationally renowned artists per group. Each group interviewed the applicant (approximately 40-minutes) and subsequently discussed their assessments within and across jury groups. During this round, both digital artworks and those physically brought to the interview were evaluated and then taken into consideration with the outcome of the interview. Finally, after deliberation, the two jury groups reached a consensus concerning their selection decisions.[3]

[2] To cope with the growing number of applicants, pre-selection has gradually been sub-divided into multiple rounds. For instance, in 1998, a separate pre-selection phase was added for Dutch nationals and in 2006, the applicants were further sub-categorized according to groups of nationalities. In all specifications of our estimation models, we found no effect of these changes.

[3] We distinguished among periods with different jury structures in our estimation models but found no significant effects.

3.1 | Data

Our data originates from multiple sources. First, we have comprehensive biographic information, provided by RABK, about all of the applicants. These data includes applicants' age, gender, nationality,[4] previous art education or other tertiary education, times applied, and letter of recommendation. Second, we have 11-years of jury notes taken during final selection. Third, in addition to these extensive notes, we have collected comprehensive qualitative data about the selection procedure based on non-participatory observation during all selection rounds as well as detailed interviews with the jury members and the program management. Fourth, our measure of future performance is based on artistic prestige rankings. We use ArtFacts. Net, a web-based platform established in 2001 that ranks contemporary visual artists based upon their annual exhibitions at galleries and museums worldwide. The platform contains exhibition history information for approximately 550,000 visual artists based on 750,000 exhibitions provided by over 35,000 galleries, museums, and other venues worldwide. ArtFacts. Net data has been widely used as a measure of performance in the management science literature (Ertug et al., 2016) and cultural sociology (Velthuis, 2013; Yogev & Grund, 2012). We use individual rankings[5] based on 2016 public art exhibitions and artists' 7-year average ranking for 2010–2016. As mentioned earlier, we also measure performance by auction sales. Notably, previous research has shown that contemporary visual artists – of which our data is comprised - are usually not active on the secondary or auction market (Prinz et al., 2015; Velthuis, 2013). This is because artists first gain legitimacy and build their reputation on the primary art market (Kackovic & Wijnberg, 2020) by exhibiting their artworks at museums and art galleries before a select few enter the secondary or auction market. Nevertheless, and exactly because RABK is a highly prestigious and renown art residency program, we investigate sales made at auction as a supplementary metric. We collect these data from Artnet.com, which is an art market website that was established in 1989 and provides detailed information about more than 9 million public art auction results from 1,600 international auction houses (Artnet.com annual report, 2014). For details, see Section 6).

4 | PRIMARY ANALYSIS OF PRE-SELECTION AND FINAL SELECTION

Table 1 presents the descriptive statistics of the pre-selection and the final selection rounds (for detailed variable descriptions and operationalizations, see Table A1 in the online Appendix). We also present summary statistics describing those applicants with and without interview notes. During pre-selection, 8.2% of the 8,557 applicants are invited for an interview. Of those invited, 21.5% have a letter of recommendation and 12.4% applied more than twice compared to those applicants not invited for an interview (14.8% and 9.1% respectively). On artistic prestige rankings, our focal performance metric, the interviewed applicants score better in relation to those who are not interviewed. For instance, almost 70% are mentioned in ArtFacts. Net and have a substantially better (lower) artistic prestige score compared to those not interviewed. More precisely, the percentage difference between the two groups is 35.3% based on the ArtFacts. Net 7-year mean rank.

Once passing the pre-selection threshold, during final selection the likelihood of being accepted to the program is 43%. A greater share of the accepted applicants (23.7%) have a letter of recommendation compared to those not admitted (15.3%) and of those admitted, applying more than once is substantially less common (3.3%) compared to those who were rejected (19.2%). Additionally, admitted applicants have a better performance status compared to the other applicants. For instance, almost 75% are mentioned in ArtFacts. Net and have a substantially better (lower)

[4]According to the head of the residency program, the selection committee considers applicants as being Dutch if they are based - and professionally (artistically) embedded - in the Netherlands.

[5]Artists are allocated points based on whether or not the exhibition was solo or group at private galleries or public institutions or biennales and other regular exhibitions, and the geographic location, with art centers such as London and New York City receiving more points than less well known ones. The artists with the highest number of points are given the lowest rank, e.g., the top five ranked artists in 2016 were: Andy Warhol, Pablo Picasso, Bruce Nauman, Gerhard Richter, and Joseph Beuys.

TABLE 1 Descriptive statistics of applicant groups

		Pre-selection		Final selection		Jury notes	
	All applicants	Interviewed applicants	Not interviewed applicants	Accepted applicants	Rejected Applicants	Yes	No
Number of observations	8,557	702	7,855	300	402	355	347
% accepted	3.5	42.3	0.0	100	0.0	44.2	40.3
% female	50.4	47.0	50.7	45.0	50.0	51.5	42.4
% letter of recommendation	7.5	21.5	14.8	23.7	15.3	23.7	19.3
% applied > 2 times	9.4	12.4	9.1	3.3	19.2	9.6	12.4
% autodidact	0.3	3.8	0.02	4.7	3.2	3.1	4.6
% unknown	91.6	0.0	99.0	0.0	0.0	0.0	0.0
Art discipline							
% sculpture	29.5	27.5	33.7	29.7	33.3	33.5	21.0
% video/film	14.9	24.2	16.0	27.7	16.3	26.8	21.6
% painting	26.6	29.3	29.9	23.3	30.2	23.1	35.4
% drawing/graphic	6.7	7.5	7.6	7.3	7.6	6.8	8.6
% photography	11.2	11.4	12.7	12.0	12.6	9.9	13.3
Artistic prestige rankings							
% mentioned in ArtFacts. Net	54.3	68.8	52.9	74.3	53.5	69.0	68.6
% ArtFacts. Net 7-year average rank	33.5	52.3	31.8	60.3	46.3	50.9	53.6
% ArtFacts. Net 2016 rank	33.1	51.9	31.5	60.0	46.0	50.7	53.3
ArtFacts.Net 7-year average rank	34,820	25,350	36,210	22,606	28,021	24,545	26,135
ArtFacts.Net 2016 rank	36,063	28,164	37,229	26,375	29,904	27,169	29,132

Note: See Section 4 for details.

ranking than those who were rejected. The percentage difference between these two groups, based on the ArtFacts. Net 7-year average rank is 21.3%, which is lower (better) ranking compared to pre-selection.

4.1 | Artistic prestige rankings for pre-selection and selection

We have complete artistic prestige ranking data for all applicants. We now take a closer look at the distribution of these rankings for the pre-selection round (see Figure 1) and the final selection round (see Figure 2).

The boxplot distribution of ArtFacts. Net rankings in Figure 1 illustrates the rank scores for those applicants invited for an interview (d = 1) and those who were not (d = 0). The horizontal lines, from top to bottom, show the maximum, 3rd quartile, median, 1st quartile, and minimum score, respectively. Individuals without Artfacts. Net scores are excluded. A dot represents an individual outlier, defined as a value greater than 1.5 interquartile ranges away from the 25th or 75th percentile. Both in 2016 and in the 7-year average ArtFacts. Net rankings, the

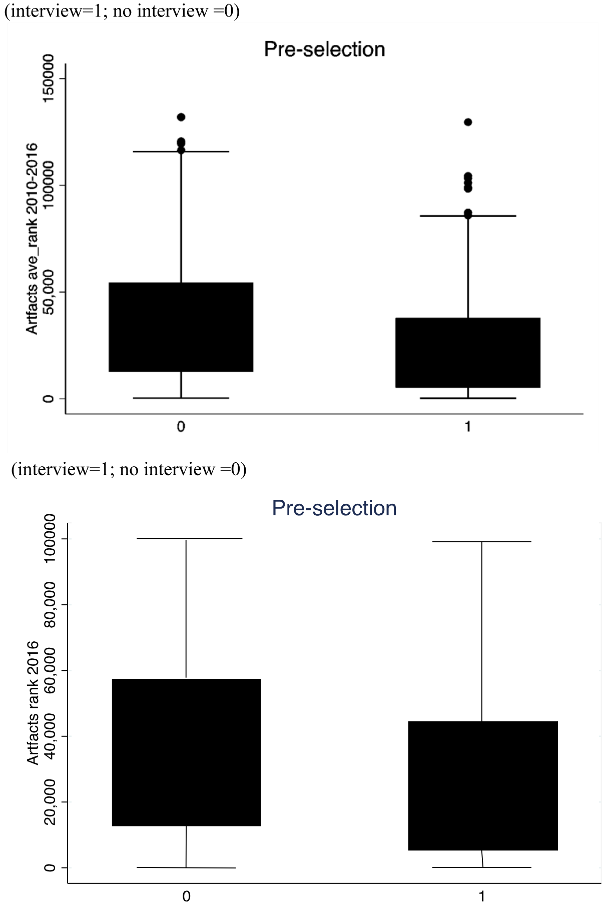
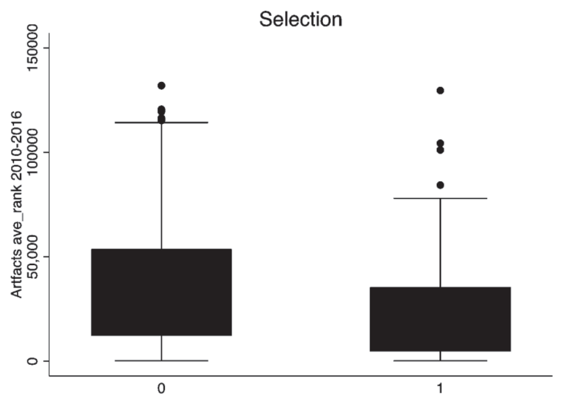


FIGURE 1 (top) Pre-selection: ArtFacts.Net 7-year average rank. (bottom) Pre-selection: ArtFacts.Net 2016 rank. The maximum, 3rd quartile, median, 1st quartile, and minimum score is indicated by a horizontal line. A dot represents an outlier, defined as a value greater than 1.5 inter-quartile ranges away from the 25th or 75th percentile [Colour figure can be viewed at wileyonlinelibrary.com]

(accepted=1; rejected =0)



(accepted=1; rejected =0)

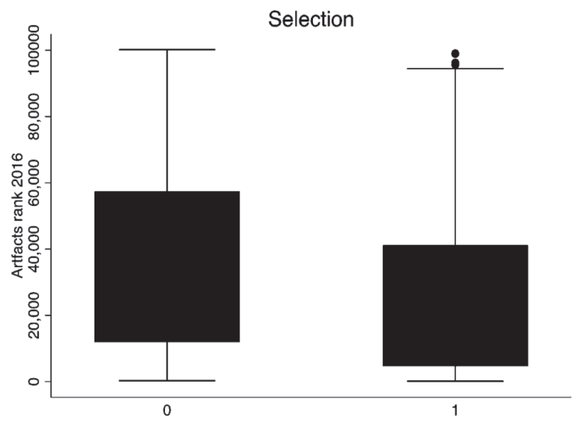


FIGURE 2 (top) Final selection: ArtFacts.Net 7-year average rank. (bottom) Final selection: ArtFacts.Net 2016 rank. The maximum, 3rd quartile, median, 1st quartile, and minimum score is indicated by a horizontal line. A dot represents an outlier, defined as a value greater than 1.5 inter-quartile ranges away from the 25th or 75th percentile

TABLE 2 Distribution of future artistic prestige rankings related to being invited to an interview in the final selection round

	ArtFacts. Net 7-year average rank		ArtFacts. Net 2016 rank	
	Low (best) rank	High (worst) rank	Low (best) rank	High (worst) rank
N	Percentage accepted among the highest and lowest ranked applicants			
5	80%	0%	80%	0%
10	50%	10%	50%	10%
25	44%	4%	44%	8%
50	28%	6%	28%	8%
100	23%	8%	23%	7%

Note: See Section 4 for details.

distribution for the interviewed applicants is shifted towards better performance compared to those not invited for an interview. The former distribution is more compressed, reaches closer to top performance and has a more dispersed bottom quartile. The Boxplots in Figure 2 show the distributions of ArtFacts. Net rankings for accepted applicants (d = 1) and the rejected ones (d = 0). These are not strikingly different than those for pre-selection in Figure 1. In both Figure 1 and Figure 2, greater heterogeneity among the interviewed and accepted applicants is underscored.

Next, we provide an additional perspective on the final selection round by considering the distribution of future artistic prestige rankings related to being invited for an interview in the final selection round (see Table 2). As mentioned earlier, during our 11-year observation period 300 of the 8,557 applicants are admitted to the program. Half of the applicants noted in the top-10 on ArtFacts. Net artistic prestige rankings have been accepted, and a little less than a quarter among the top-100. The share of the accepted applicants declines with N, the defined length of the top. While the acceptance rate at the bottom is less sensitive to the defined length as it is consistently low or zero. Seemingly, weeding out applicants with less potential is easier than recognizing future top performers. As evidenced from the bottom-100 of the artistic prestige rankings, more than 9 out of 10 applicants have been rejected.

4.2 | Admitting “better” applicants

In Table 3, we consider the probability of an applicant being invited for an interview and the probability of being admitted to the program in relation to their future artistic prestige. However, we disregard the reverse effect that the program may have on later success. Our argument to do so is that we assume performance is exogenous and that the objective of the selection procedure is to admit the “better” artists. We surmise that identifying the potential of the applicant the jury members aim for is reflected in ArtFacts. Net rankings. We argue that regressing admittance on performance reveals the success of the selection procedures. Of course, this holds only under the premise that ArtFacts. Net measures the potential of the applicants the jury aims to discover and that performance is independent of the program.

As older cohorts of applicants have had more time to realize their potential (Marconi, 2018), we control for the years elapsed since leaving the program by measuring performance as the residual from a regression on career age and its square. We also indicate with a dummy variable whether the individual has an Artfacts. Net ranking. The results in Table 3 indicate that realized future artistic success, i.e., a better (lower) ArtFacts. Net ranking, has a significant effect on the probability of being invited for an interview. We conclude that artists who turn out to perform better had a greater chance of continuing to the final selection round and being accepted to the program.

From the analyses presented here, we infer that these highly selective procedures leads to admission of applicants who later score substantially better (lower) on artistic prestige rankings. The largest part of the gain in the average score of admitted applicants compared to all applicants is due to the pre-selection round.

4.3 | Jury notes

The descriptive statistics in Table 1 suggest that applicants without jury notes are the marginally weaker group. In other words, they are accepted at a slightly lower frequency, and a smaller share has a letter of recommendation while a larger share does not have formal art school training. Additionally, slightly more individuals from this group have not been mentioned by ArtFacts. Net, and of those who are ranked, their ranking is less favorable than those applicants with jury notes.

We have interview notes for 52% of the accepted of applicants (157) and 49% of the rejected ones (198). Table 4 presents the estimates from Probit regressions on the probability of having jury notes. We do not find a significant effect on the acceptance decision. In sum, these results confirm that the availability of jury notes is random, and therefore, not selective in relation to individual quality or qualification.

TABLE 3 Probability of an interview and selection to the program related to future artistic prestige rankings

	Outcome variable			*Outcome variable*		
	Probability of an interview			**Probability of selection to the program**		
ArtFacts.Net 7-year average rank						
Residual (log) average rank from 2010–2016	−0.164** −(0.015)			−0.125*** (0.032)		
Ranked 2010–2016 in Artfacts. Net rank	−1.984** (0.221)			−1.537** (0.496)		
ArtFacts.Net 2016 rank						
Residual (log) 2016 rank		−0.157** (0.015)			−0.113*** (0.032)	
Ranked 2016 in Artfacts.Net		−1.880** (0.218)			−1.359** (0.487)	
Not ranked in Artfacts.Net			−0.428** (0.040)			−0.354*** (0.096)
N	8557	8557	8557	702	702	702
Pseudo R^2	0.049	0.048	0.024	0.030	0.028	0.014

Note: The estimation methods used are Probit regressions. See Section 4 for details. Residuals indicate the predicted value controlling for career age and career age squared. The presented results are coefficient estimates with robust standard errors in parentheses.
***Significance level at 1%
**Significance level at 5%.
*Significance level at 10%.

5 | LATENT STRUCTURE OF CRITERION VARIABLES

As mentioned earlier, the selection procedure for this program is not institutionalized or formally structured. While there is a non-structured discussion of an applicant's qualities as assessed by the jury members and final selection is made based on jurors' subjective assessments, there is no formal voting or explicit scoring on defined criteria. That, however, does not rule out that an underlying structure reflected in an implicit set of selection criteria may guide the decision process. To further investigate this possibility, non-participatory observation and interviews with jury members were used to gain a finer grained understanding of the latent structure of the selection criterion variables.

5.1 | Qualitative analysis

Over a four month period, one of the authors observed both the pre-selection and final selection rounds. After each round, semi-structured interviews were held with each jury member. The interviews were designed so that each juror could rank the importance of the criteria that were observed during the focal rounds. In addition to the jury interviews, multiple in-depth interviews were conducted with the head of the residency program as well as the director of RABK. Both independently ranked the criteria deduced from the non-participatory observation, and evaluated

TABLE 4 Probability of having interview notes related to admission and future artistic prestige rankings

Outcome variable					
Availability of interview notes					
	Interview notes (1)	**Interview notes (2)**	**Interview notes (3)**	**Interview notes (4)**	**Interview notes (5)**
Accepted	−0.011 (0.117)	−0.001 (0.119)	0.001 (0.119)	0.009 (0.119)	0.004 (0.119)
Letter of recommendation	0.067 (0.136)	0.084 (0.136)	0.084 (0.137)	0.076 (0.136)	0.080 (0.137)
Sculpture	0.615** (0.191)	0.630** (0.192)	0.620** (0.191)	0.617** (0.191)	0.619** (0.195)
Video	0.410** (0.198)	0.399** (0.199)	0.394** (0.199)	0.413** (0.199)	0.390** (0.199)
Painting	−0.094 (0.191)	−0.088 (0.192)	−0.091 (0.192)	−0.103 (0.192)	−0.081 (0.193)
Drawing	0.092 (0.253)	0.095 (0.253)	0.091 (0.253)	0.093 (0.253)	0.085 (0.254)
Female	0.220** (0.111)	0.220** (0.111)	0.221** (0.111)	0.221** (0.111)	0.218** (0.111)
Nationality	−0.054 (0.124)	−0.057 (0.125)	−0.057 (0.125)	−0.071 (0.125)	−0.061 (0.126)
Age at application	0.025 (0.016)	0.024 (0.016)	0.025 (0.016)	0.024 (0.016)	0.025 (0.016)
Times applied	−0.098 (0.071)	−0.090 (0.069)	−0.091 (0.069)	−0.090 (0.072)	−0.092 (0.069)
Autodidact	0.027 (0.298)	0.001 (0.298)	−0.006 (0.299)	0.002 (0.297)	−0.004 (0.297)
Residual (log) average rank from 2010–2016		−0.076 (0.059)			0.042 (0.253)
Ranked during 2010–2016 in Artfacts. Net rank		−1.285 (0.924)			0.472 (3.587)
Residual (log) 2016 rank			−0.075 (0.054)		−0.120 (0.229)
Ranked in 2016 Artfacts.Net			−1.264 (0.842)		−1.766 (3.252)
Not ranked in Artfacts.Net				0.112 (0.114)	
N	702	702	702	702	702
Pseudo R^2	0.260	0.263	0.263	0.261	0.263
Year fixed effects	Yes	Yes	Yes	Yes	Yes

Note: The estimation methods used are Probit regressions. See Section 4 for details. Residuals indicate the predicted value controlling for career age and career age squared. The presented results are coefficient estimates with robust standard errors in parentheses.

***Significance level at 1%.

**Significance level at 5%.

*Significance level at 10%.

TABLE 5 List of implicit selection criteria

Selection criteria:	Definitions:
Content	refers to the development or relevant progress of the artist over time.
Context	refers to the artists' understanding of art-historical, political, and social references in relation to the content of their own work.
Craft	refers to the technical ability to make the artwork.
Creativity	refers to the innovative aspects and novel contribution of the artwork.
Expertise	refers to the artist's art education or education in other academic disciplines, e.g., biology, architecture.
Critique	refers to how open the artist is to criticism and change.
Communication	refers to the artist's ability to communicate about the artworks.
Fit	refers to the stage of the artist's career and the fit within the RABK community.
Collaboration	refers to if the jury member(s) would like to have the artist as a colleague if they are challenged and intrigued to have ‘studio visits’ with the applicant.
Personality	refers to if the jury member(s) personally liked the applicant.
Final assessment	refers to jury members' overall assessment of the applicant; this score does not follow mechanically from the scores on the other criteria, it is, inter alia, based on individual juror's subjective assessments.

the jury members' aggregate ranking per criterion. There was complete agreement between the jury members and RABK management. Table 5 provides the listing of the uncovered implicit selection criteria that were used during final selection (for additional details, see Table A4 in the online Appendix).

We are confident that our representation of the selection procedure is valid for the entire observation period. The head of the residency program has been involved in this process since 1995, and the upper management team remained relatively stable. Although the director, who had held that position for 28 years, retired in 2010, his successor was familiar with the art residency program, as she had served on the management team since 1995. Additionally, there has been a high degree of continuity in the jury's composition, with 21 out of 30 members serving two or more years.

5.2 | Analysis of interview notes

The jury notes from the interviews conducted during the final selection round have been transcribed verbatim. Using the latent structure framework of the selection criterion variables uncovered during the qualitative analysis, three raters scored the notes on each of the criteria.[6]

The raters read the interview notes and then assigned a value to each criterion. A value of 1 represented a negative jury assessment, a value of 2 was neutral, and a value of 3 was positive. If there was no explicit comment in the notes on a particular criterion - the applicant scored neither clearly positive or negative - a value of 2 was given. We included a final assessment score, where a value of 1 represented an overall negative evaluation, and a value of 3 represented an overall positive evaluation.[7] A two-way random effects ICC model[8] was used to measure the

[6]The raters included one of the authors and two MSc in Business Administration students. Before evaluating and scoring the notes, the students received extensive coding training.

[7]In 8 observations, a final jury assessment score was missing; all these applicants were rejected. Excluding these observations from regression analyses proved immaterial.

[8]Intraclass correlation coefficient (ICC) is used to measure the consistency or homogeneity of measurements given to the same target by different raters (Shrout & Fleiss, 1979).

consistency of agreement. The result showed that the consistency of absolute agreement between all three raters per criterion was satisfactory.[9]

On their website, RABK describes the application procedure: "Artistic quality, possibilities for development and the place in applicants' career are considered." We have also created an implementation of these key variables based on the latent selection criteria.[10] *Artistic quality* is the average score of craft, creativity, context, and communication. While craft and creativity are obvious components of artistic quality, the other selection criteria such as context, and communication reflect artists' understanding of the art world and the ability to communicate about their work. In other words, it is a measure of the awareness of the artist's position within the wider artistic environment. *Development potential* is the average score of content and critique. By definition, content refers to development, and openness to critique is helpful to enhance development. Finally, *group composition* is measured by the average score of fit, personality, collaboration, and expertise. The definition of fit refers not only to the target age group of between 27 to 36 years old and independent art practice experience but also explicitly to the suitability within the RABK community. For example, a pleasant personality may smooth group interactions, and the same holds for an artist you would like to have as a colleague. Expertise refers to education, specifically in combination with other fields. Thus, a broader perspective may bring new insights to the group.

5.3 | Quantitative analysis

In Table 6, we test the existence of the latent structure by estimating Probit regressions on biographical data and then adding the implicit selection criteria uncovered during our qualitative analysis. Regression model (1) includes biographical data on the full sample of interviewed applicants in the final selection round. The pseudo R^2 is 0.102. We can only analyze the existence of the latent structure if jury notes are available, hence we repeat the analysis in model (2) for the sub-sample of applicants with notes ($n = 355$). The pseudo R^2 is 0.129. In model (3) we add selection criteria uncovered during the qualitative analysis. We find that two criteria – creativity and fit – have a significant and positive effect on the probability of acceptance to the program. The pseudo R^2 doubles to 0.237. In model (4) we add the residual of final assessment. As noted earlier, this variable is not just a summary of the separate criteria scores. Estimation shows that six out of ten criteria have a significant effect on final assessment, but substantial residual variation remains. Therefore, we use the residual of final assessment, after controlling for the separate criteria, to account for unobservable selection considerations. We show that the effects of biographical data, selection criteria and residual final assessment are essentially independent. In model (5) we include both the criteria from the jury notes and the residual of final assessments. The pseudo R^2 increases to 0.345. Finally, as model (6) shows, the condensation of the jury criteria into three criteria leads to 3 variables with significant effects. It is interesting to note that the 3 aggregate criteria have about equal weight, taking into account that development potential is based on two aggregated criteria instead of four. The pseudo R^2 is slightly reduced to 0.308.

In summary, we have uncovered a latent structure of criterion variables that helps explain the jury's admission decisions. Compared to using biographical data only, the pseudo-R^2 increased from 0.129 to 0.237 when we added the latent criterion variables. Condensing the variables to three composites that reflect the stated goals of RABK only slightly reduces the pseudo-R^2, thus confirming that these latent variables reflect a meaningful structure.

[9]According to Shrout and Fleiss (1979) an ICC score of 0.70 is considered acceptable. Across all our criteria, the score was 0.87.
[10]We verified the factor structure of our variables with confirmatory factor analysis (CFA).

TABLE 6 Probability of admission to the program in the final selection round

Outcome variable						
Selection to the program						
	All interviewed applicants (1)	**Interviewed applicants with jury notes (2)**	**Interviewed applicants with jury notes (3)**	**Interviewed applicants with jury notes (4)**	**Interviewed applicants with jury notes (5)**	**Interviewed applicants with jury notes (6)**
Age	−0.052 (0.088)	−0.111 (0.181)	0.013 (0.169)	−0.243 (0.249)	−0.110 (0.258)	−0.079 (0.226)
Female	−0.148 (0.102)	−0.099 (0.146)	−0.122 (0.155)	−0.210 (0.156)	−0.222 (0.168)	−0.217 (0.163)
Nationality	0.255** (0.115)	0.324* (0.172)	0.416** (0.187)	0.482*** (0.185)	0.598*** (0.205)	0.578*** (0.198)
Autodidact	0.285 (0.267)	0.059 (0.429)	0.386 (0.493)	0.257 (0.446)	0.605 (0.516)	0.685 (0.490)
Letter	0.213* (0.127)	0.537*** (0.179)	0.556*** (0.191)	0.512*** (0.189)	0.519** (0.204)	0.547*** (0.198)
Times applied	−0.540*** (0.078)	−0.445*** (0.112)	−0.463*** (0.115)	−0.414*** (0.120)	−0.421*** (0.124)	−0.423*** (0.122)
Sculpture	0.044 (0.177)	−0.035 (0.268)	−0.083 (0.287)	−0.085 (0.281)	−0.093 (0.302)	−0.179 (0.289)
Video	0.164 (0.182)	−0.043 (0.273)	−0.052 (0.297)	−0.116 (0.286)	−0.094 (0.314)	−0.182 (0.298)
Painting	−0.204 (0.178)	−0.371 (0.281)	−0.471 (0.306)	−0.363 (0.293)	−0.455 (0.319)	−0.504* (0.319)
Drawing	−0.139 (0.234)	−0.508 (0.374)	−0.761* (0.400)	−0.430 (0.382)	−0.638 (0.409)	−0.628 (0.395)
Content			0.134 (0.099)		0.139 (0.106)	
Context			0.157 (0.141)		0.138 (0.152)	
Craft			0.098 (0.117)		0.092 (0.125)	
Creativity			0.436*** (0.100)		0.488*** (0.106)	
Experience			0.076 (0.170)		0.040 (0.179)	
Critique			0.205 (0.154)		0.218 (0.163)	
Communication			0.035 (0.123)		0.029 (0.130)	
Fit			0.445*** (0.121)		0.424*** (0.128)	
Collaboration			0.229 (0.228)		0.223 (0.242)	
Personality			0.041 (0.142)		0.047 (0.148)	

(Continues)

TABLE 6 (Continued)

Outcome variable						
Selection to the program						
	All interviewed applicants (1)	**Interviewed applicants with jury notes (2)**	**Interviewed applicants with jury notes (3)**	**Interviewed applicants with jury notes (4)**	**Interviewed applicants with jury notes (5)**	**Interviewed applicants with jury notes (6)**
Artistic quality						0.834*** (0.229)
Development potential						0.401*** (0.173)
Group composition						0.720*** (0.320)
Residual final assessment				0.641*** (0.091)	0.652*** (0.094)	0.635*** (0.092)
N	702	355	355	355	355	355
Pseudo R^2	0.102	0.129	0.237	0.239	0.345	0.308

Note: The estimation methods used are Probit regressions. See Section V for details. The residual final assessment variable measures unobservable selection considerations after controlling for the observed selection criteria. See Section 5 for details. The presented results are coefficient estimates with robust standard errors in parentheses.
***Significance level at 1%.
**Significance level at 5%.
*Significance level at 10%.

6 | PREDICTING PERFORMANCE

We investigate the extent to which the selection procedure predicts artists' future performance by estimating a switching regression model including accepted and the rejected applicants and correcting for selectivity bias. Switching regressions are similar to two-step selection models (Heckman, 1974, 1976), i.e., to get consistent estimates, the mean zero restriction on the error term, conditional on the explanatory variables, is restored by including an estimate of the selection bias. This means that these models are efficient as long as the functional relationship between the outcome equation and the selection equation is correctly specified.

Newey (2009) suggests a method to estimate the Heckman selectivity model semi-parametrically. The switching regression models are estimated in a sequential procedure that first estimates a binary, say 0 or 1, choice equation for selection and then estimates two linear regressions for performances for the choice 0 and for the choice 1. We denote the selection process in the following fashion:

$$I_i^* = X_i'\beta + \varepsilon_i \tag{1}$$

where the individual i is selected into option 1 if $I_i^* \geq 0(\text{i.e.} I_i = 1)$ and option 0 if $I_i^* < 0(\text{i.e.} I_i = 0)$. We denote the performance equations as:

$$y_{ij} = Z_i'\alpha_j + v_{ji} \tag{2}$$

This performance is only observed for the individual's selection outcome. If ε_i correlates with v_{ji}, an OLS estimation will yield biased estimation results. To correct for this, we need to correct for $E(y_{ij}|I_i = j)$. Under the assumption of

TABLE 7 Probability of admission and future artistic prestige rankings

Selection equation		Performance equations			
Outcome variable		*Outcome variable*			
		Artfacts.Net 7-year average rank		Artfacts.Net 2016 rank	
Selection to the program		*Accepted applicants*	*Rejected applicants*	*Accepted applicants*	*Rejected applicants*
Female	−0.481*** (0.182)	0.426 (1.313)	−0.433 (1.102)	0.145 (1.296)	−0.420 (1.152)
Nationality	1.442*** (0.325)				
Age at application	−0.017 (0.024)				
Career age		0.112 (0.213)	0.364* (0.194)	0.140 (0.226)	0.385* (0.205)
Times applied	−0.790*** (0.226)	−2.214 (1.976)	−0.646 (1.111)	−2.314 (1.840)	−0.677 (1.076)
Sculpture	0.072 (0.286)	−0.394 (2.270)	−2.178 (2.667)	−0.670 (2.102)	−2.380 (2.535)
Video	0.093 (0.304)	−1.767 (2.338)	−1.029 (2.588)	−1.973 (2.179)	−1.114 (2.498)
Painting	−0.160 (0.337)	−0.540 (2.435)	−0.156 (2.736)	−0.324 (2.397)	−0.246 (2.470)
Drawing	−0.755* (0.457)	−1.994 (3.272)	−3.718 (3.949)	−2.270 (3.219)	−3.706 (3.496)
Letter	0.298 (0.235)	−1.594 (1.703)	−0.733 (1.700)	−1.340 (1.745)	−0.681 (1.623)
Autodidact	0.523 (0.624)	2.507 (4.025)	−1.441 (4.811)	5.262 (3.934)	−1.317 (4.609)
Artistic quality	1.310*** (0.249)	−0.136 (2.022)	1.631 (2.235)	0.096 (2.171)	1.610 (2.063)
Development potential	0.573*** (0.205)	−0.201 (1.437)	−0.981 (1.321)	−0.035 (1.429)	−0.909 (1.331)
Group composition	0.768*** (0.296)	1.261 (2.350)	−0.793 (2.272)	1.349 (2.289)	−0.729 (2.112)
Residual final assessment	0.925*** (0.194)	0.098 (0.768)	−0.450 (0.746)	0.239 (0.810)	−0.456 (0.750)
N	355	355	355	355	355

Note: The estimation methods used are switching regressions, similar to two-step selection models (Heckman, 1974, 1976). The estimations are conducted in a sequential procedure that first estimates a binary choice equation for selection and then estimates two linear regressions for performance for the choice *accepted applicants* and for the choice *rejected applicants*. See Section 6 for details. The residual final assessment variable measures unobservable selection considerations after controlling for the observed selection criteria. See Section 5 for details. The presented results are coefficient estimates with robust standard errors in parentheses.

***Significance level at 1%.

**Significance level at 5%.

*Significance level at 10%.

normal distributed error terms this expectation is equal to the inverse Mill's ratio (Heckman, 1974). Without this assumption, it can be stated that:

$$
\begin{aligned}
E(y_{1i}|I_i=1) &= Z_i'\alpha_1 + E\left(v_{1i}\middle|\varepsilon_i \geq -X_i'\beta\right) = Z_i'\alpha_1 + g_1\left(X_i'\beta\right) \\
E(y_{0i}|I_i=0) &= Z_i'\alpha_0 + E\left(v_{0i}\middle|\varepsilon_i \geq -X_i'\beta\right) = Z_i'\alpha_0 + g_0\left(X_i'\beta\right)
\end{aligned}
\tag{3}
$$

Newey (2009) proposes to approximate the unknown functions $g_j\left(X_i'\beta\right)$ by a polynomial $\sum_{k-1}^{K}\eta_{jk}\tau(X_i\beta)^{k-1}$ where $\tau(\cdot)=2\Phi(\cdot)-1$ and Φ the standard normal cumulative distribution function. Newey proves that by adding the approximation of $g_j\left(X_i'\beta\right)$ to the linear equation, consistent estimation results can be found by employing OLS if the order of the polynomial is increased with N.

To get full semi-parametric estimation of the model, the selection equation is estimated without assuming a specific distribution of the error term. We use the semi-nonparametric estimation method proposed by Gallant and Nychka (1987). This estimation also relies on a polynomial approximation. In practice, the order of the polynomial in both the Newey, and Gallant and Nychka methods has to be chosen rather low. We opt for a polynomial of order 3 in both cases.[11] Because we employ a two-step estimation procedure, the standard errors of the estimated parameters need to be bootstrapped. We use a nonparametric bootstrap with 250 replications.

We have 355 observations for both performance and selection equations for accepted and rejected applicants. We specify a performance equation for prestige rankings, measured by ArtFacts.Net. We then consider each artist's prestige rankings in 2016 and the 7-year mean rank. We assigned those individuals who were not scored in 2016 a score of 25 (log), which is artificially high. The range of ArtFacts. Net scores expressed in the natural log is 5.01 (for the best performer), 11.82 (for the worst), and 25 (for those not ranked).

In the performance equation, we estimate the relationship between artistic prestige rank and the explanatory variables, i.e., biographical data and selection criteria. We also specify the variable career age, which is measured as 2015 minus the year invited to the interview. We include this variable to control for experience, i.e., artists with longer careers had opportunities to sell more artworks and participate in more exhibitions compared to artists with shorter career trajectories. In the selection equation, we estimate the binary decision of acceptance and rejection, using the same explanatory variables as in the performance equations. To enhance identification, we use the variable "nationality" only in the selection equation and exclude it from the performance equation, on the argument that while RABK has preferences for the country of origin,[12] the contemporary art world operates internationally (see e.g., Crane, 2009; Khaire, 2015; Velthuis & Curioni, 2015). We have included the aggregate specification for the selection criteria, in three variables, as a compact representation of the jury's translation of RABK targets for admissions.

Table 7 shows that the selection equation only differs in detail from the estimation as a single regression equation in Table 6. Nationality, times applied, the three aggregated criteria, and the residual of final assessment have a significant effect. Gender has a significant effect (females have a smaller chance of acceptance), the letter of recommendation is no longer significant, and the relative position of disciplines has slightly changed. The relative weights of the criteria have changed slightly, but the ranking - artistic quality, group composition and development potential- remains the same. These results show that the variables used in the jury decisions do not have a significant effect on performance, either for the accepted or for the rejected applicants. The one exception being the career age of the rejected applicants in the ArtFacts. Net rankings.

We have formally tested equality of the coefficients in the performance equations for the accepted and the rejected applicants (rank 2016, and average 7-year rank), in three specifications: without correcting for selectivity (just the linear part of the model), with selectivity corrections for the performance of the accepted and the

[11]We experimented with higher-order polynomials. The estimation results are very similar, however the precision decreases considerably.

[12]RABK aims for a substantial share of Dutch nationals but is not explicit about other countries.

rejected but equal intercepts and with selectivity corrections and unconstrained intercepts. In none of these specifications were differences statistically significant at 5% (for details, see Table A2 in the online Appendix). This indicates that there are no separate “production functions” for artistic success among the accepted and the rejected.

We have also tested the significance of the terms for selectivity correction, as a test on the correlation between errors in performance and selection equations (for details, see Table A3 in the online Appendix). The correction term is significant (at 10%) for the rank of accepted applicants but not significant for the rejected. Thus, unobservables in the selection equation correlate with the unobservables in the 2016 performance equation for admitted applicants, but such a relationship does not hold for the performance of rejected applicants. The results show that there is some quality component affecting the later performance of the admitted applicants that the jury acknowledges in the selection that the focal variables do not explicitly measure. We conclude that the switching regression model is only weakly superior to the single equation regression model. Namely, performance equations are not statistically different; the need for selectivity correction is recognized only for accepted applicants and not for the rejected ones. This implies that the single regression equation model estimated on accepted applicants suffers mildly from a selectivity bias that emerges from heterogeneity among the applicants.

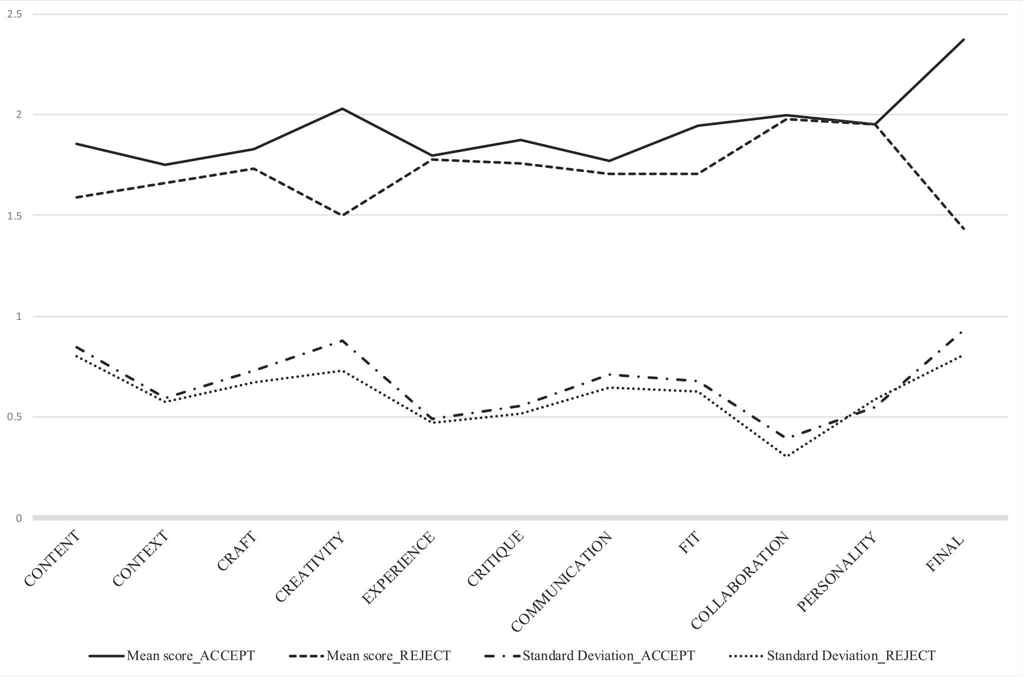
With insignificant effects of the selection criteria in the performance equations, the major difference among the accepted and the rejected applicants is the difference in qualifications at entry, as a consequence of selection. Figure 3a and Figure 3b show that the accepted applicants score higher than the rejected ones on all jury criteria, both for the 10 separate criteria and the 3 aggregate criteria we have constructed.

6.1 | Robustness Tests

We check the robustness of our estimations in two distinct ways. First, with Probit regressions to estimate the probability of admission using alternative operationalizations of our predictor variables. Second, switching regressions to estimate future sales, our alternative outcome variable. We start with three new operationalizations of our ten selection criteria (predictor variables). First, we create a dichotomous variable measuring if the jury commented on a particular criterion during the evaluation of the applicant; if yes, a value of 1 is given, otherwise a 0. However, we do not consider the valence of the comment. Second, we create a variable indicating the maximum score the applicant received per criterion, as determined by the highest score given by any of the three raters. In this operationalization, we consider the highest valence. Third, we use only the score of one rater, namely, one of the authors (for details, see Section 5). We present our findings in Table 8. The results show that all alternative operationalizations are robust to our main estimations presented in Table 6.

Next, we use sales at public art auctions as an alternative outcome variable. Our auction sales data, retrieved from Artnet.com, contain career sales information of artists up to and including 2015. We estimate the probability of admission and future auction sales by employing switching regressions in Table 9. For details, see Section 6. Although auction sales are not the main focus of our analysis, the percentage of any sales (13.1%), the mean number of sales (5.5), and the mean sales price of those interviewed (30 times higher) is considerably higher compared to those not invited to the interview (0.03% and 1, respectively). The core findings remain that, while the aggregate selection variables have a significant and positive effect on selection, they do not influence future auction sales. These results are robust to the artistic prestige ranking estimations presented in Table 7. To conclude, although details from the robustness checks may differ slightly concerning particular regression coefficients (although not the sign or significance), the broad conclusion is robust, namely that selection criteria have a significant and positive effect on admission but do not influence future performance.

(a)



(b)

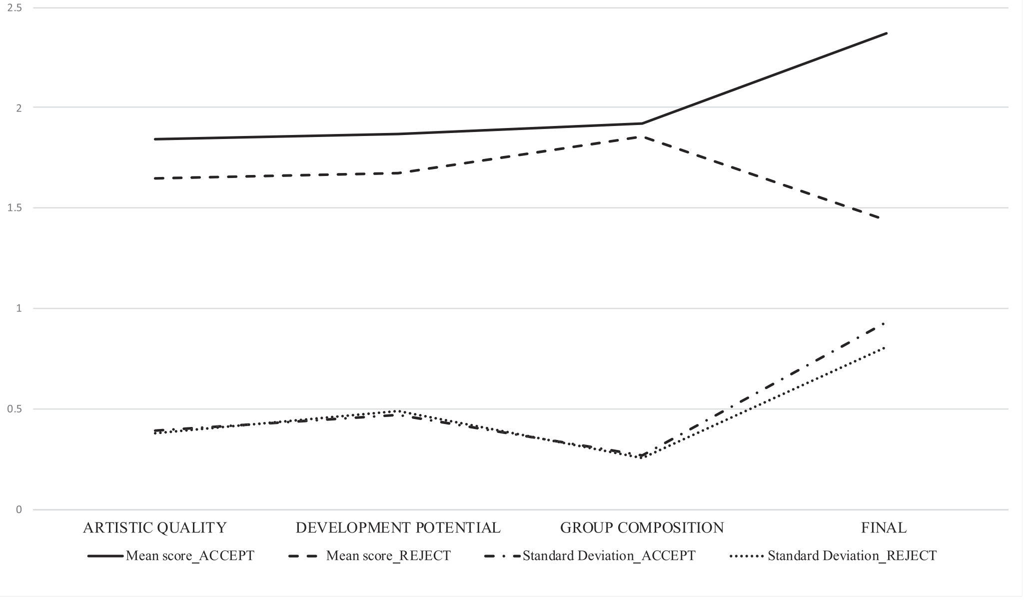


FIGURE 3 (a): Distribution of implicit criteria scores in the final selection round (b): Distribution of aggregate selection criteria in the final selection round.

TABLE 8 Probability of admission to the program in the final selection round

Outcome variable			
Selection to the program			
	Interviewed applicants with jury notes		
	Explicit mention	**Max score**	**One rater's scores**
Age	0.459 (0.022)	0.490 (0.022)	0.041 (0.022)
Female	−0.253 (0.152)	−0.269 (0.153)	−0.269 (0.153)
Nationality	0.516*** (0.182)	0.529*** (0.183)	0.518*** (0.179)
Autodidact	0.465 (0.427)	0.369 (0.427)	0.475 (0.427)
Letter	0.445** (0.186)	0.456** (0.187)	0.438** (0.185)
Times applied	−0.413*** (0.114)	−0.450*** (0.121)	−0.438*** (0.122)
Sculpture	−0.127 (0.273)	−0.162 (0.273)	−0.128 (0.270)
Video	−0.070 (0.281)	−0.134 (0.286)	−0.087 (0.280)
Painting	−0.438 (0.288)	−0.410 (0.292)	−0.360 (0.287)
Drawing	−0.390 (0.234)	−0.311 (0.377)	−0.341 (0.375)
Content	0.717 (0.242)	0.185 (0.074)	0.134 (0.072)
Context	0.184 (0.172)	0.093 (0.069)	0.161 (0.112)
Craft	0.305 (0.177)	0.080 (0.063)	0.150 (0.070)
Creativity	0.108** (0.250)	0.209*** (0.075)	0.123** (0.064)
Experience	0.126 (0.178)	0.032 (0.080)	0.066 (0.119)
Critique	0.100 (0.176)	0.011 (0.073)	−0.012 (0.090)
Communication	0.017 (0.175)	0.171 (0.117)	0.047 (0.079)
Fit	0.085** (0.160)	0.065** (0.069)	0.017** (0.075)
Collaboration	0.260 (0.241)	0.080 (0.094)	0.050 (0.119)
Personality	0.012 (0.168)	0.058 (0.086)	0.031 (0.086)
Residual final assessment	0.592*** (0.088)	0.524*** (0.086)	0.529*** (0.089)
N	355	355	355
Pseudo R^2	0.219	0.224	0.212

Note: The estimation methods used are Probit regressions. See Section 5 for details. The residual final assessment variable measures unobservable selection considerations after controlling for the observed selection criteria. See Section 5 for details. The presented results are coefficient estimates with robust standard errors in parentheses.
***Significance level at 1%.
**Significance level at 5%.
*Significance level at 10%.

TABLE 9 Probability of admission and future auction sales

Selection equation		Performance equations	
Outcome variable		*Outcome variable*	
		(log) Auction sales	
Selection to the program		*Accepted applicants*	*Rejected applicants*
Female	−0.253 (0.156)	−1.846 (1.393)	−0.609 (0.735)
Nationality	0.621*** (0.185)		
Age at application	0.002 (0.020)		
Career age		0.077 (0.173)	0.036 (0.066)
Times applied	−0.405*** (0.115)	−1.951 (1.821)	−1.123* (0.676)
Sculpture	−0.090 (0.275)	−1.820 (2.119)	−0.668 (1.136)
Video	−0.013 (0.289)	−1.614 (2.181)	−0.259 (1.135)
Painting	−0.372 (0.295)	−2.392 (2.587)	−0.309 (1.346)
Drawing	−0.532 (0.382)	−4.955 (3.561)	−0.688 (1.807)
Letter	0.396** (0.184)	1.828 (1.926)	1.380 (1.041)
Autodidact	0.605 (0.477)	2.089 (3.716)	−0.056 (2.131)
Artistic quality	0.737*** (0.148)	3.853 (2.615)	2.127 (1.466)
Development potential	0.897*** (0.231)	3.878 (3.219)	1.828 (1.913)
Group composition	0.268*** (0.260)	−1.427 (2.105)	−0.501 (1.088)
Residual final assessment	0.604*** (0.086)	2.663 (2.096)	1.375 (1.203)
N	355	355	355

Note: The estimation methods used are switching regressions, similar to two-step selection models (Heckman, 1974, 1976). The estimations are conducted in a sequential procedure that first estimates a binary choice equation for selection and then estimates two linear regressions for performance for the choice *accepted applicants* and for the choice *rejected applicants*. See Section 6 for details. The residual final assessment variable measures unobservable selection considerations after controlling for the observed selection criteria. See Section 5 for details. The presented results are coefficient estimates with robust standard errors in parentheses.

***Significance level at 1%.

**Significance level at 5%.

*Significance level at 10%.

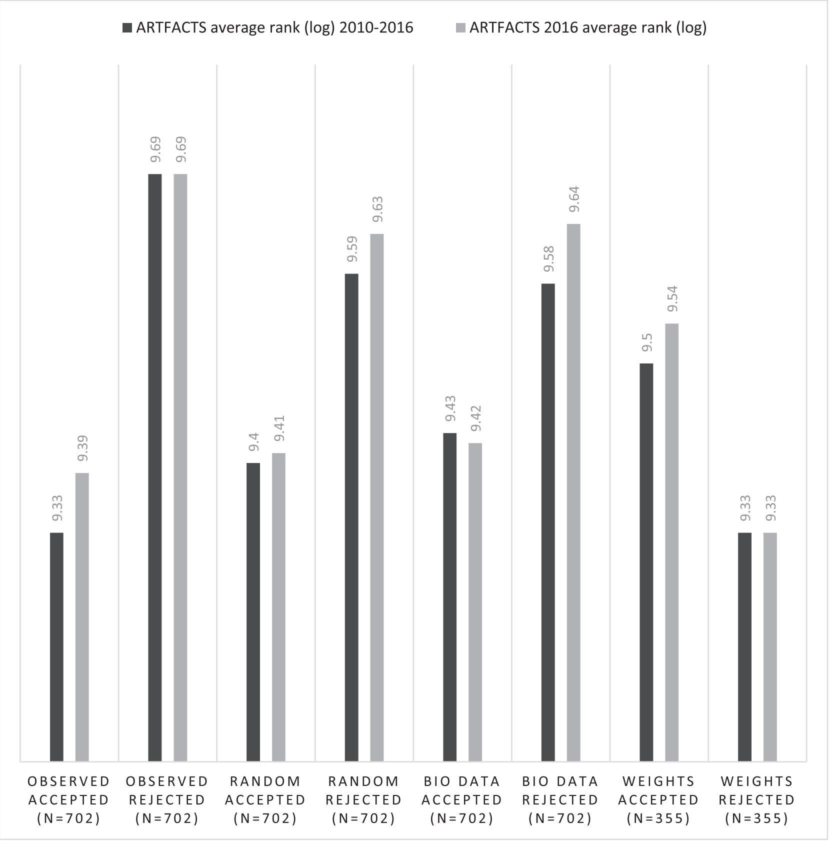


FIGURE 4 Alternative allocations compared to the observed performance in the final selection round. Alternative assignments have equal fraction of accepted as observed. Random is based on a lottery, Bio Data - as predicted by admission equation- uses demographic data, and Weights - as predicted by admission equation - uses the weighted criteria as indicated by jury members and RABK management

6.2 | Simulated alternative decision rules

Next we look at the final selection of applicants - in terms of artistic prestige rankings – and compare it to outcomes based upon alternative decision rules. We simulate a change in the composition of the sets of accepted and rejected applicants by changing the admission rules, while maintaining the proportion of applicants admitted at 43%. We do this to see if alternative selection rules would result in different compositions by performance outcomes compared to our observed population. First, using the sample of applicants invited to the interview (n = 702), we simulate selection based upon random assignment, where each applicant in our sample has a probability of acceptance of 0.43. Second, we calculate scores for the admission decision based on the coefficients estimated for the admission equation with biographical data only and admit the top 43%. Third, we replace the estimated coefficients in our main selection equation by the selection criteria weights as rated by the jury members and the head of the residency program,[13] for the sample of applicants with jury notes (n = 355), while keeping all other coefficients at their estimated value and admit the 43% with the highest scores (for details, see Table A4 in the online Appendix).

[13]Each aggregate criterion gets the sum of the weights of its constituent criteria.

In principle, alternative admission rules can change artists' status. In other words, we may admit artists who were originally rejected and reject those who were originally admitted; hence, there is a selectivity issue. However, we argue that with these data, selectivity is inconsequential because the two performance equations from the observed sample population are not statistically different from one another. Note that for each applicant, we simply use the observed performance, regardless of whether the applicant was accepted or rejected.

Our simulations show that the other alternative allocations perform worse than the observed allocation (Figure 4). Under the observed allocation, accepted applicants rank better than under the alternative allocations, and the rejected applicants rank worse. By implication, the realized allocation has the largest performance gap between accepted and rejected applicants, precisely the *raison d'étre* of selection. With standard deviations of log-rank for accepted and rejected of about 1.4, random assignment and assignment by biographical data only leads to differences in mean rank for the accepted and the rejected of no more that 6% of a standard deviation. While the assignment by stated jury weights increase the mean rank for the accepted by 10% of a standard deviation and reduce it for the rejected by 25% of a standard deviation. It is interesting that the gap in the performance of accepted applicants is largest between realized allocation and those by stated jury weights, providing additional evidence that the jury does not consciously apply a transparent selection process.

7 | DISCUSSION

The Rijksakademie van Beeldende Kunsten aims to admit exceptionally talented artists who they perceive will benefit most from the residency program and to facilitate their development to the top of the international art market. Their selection procedure consists of pre-selection, where the decision to invite an applicant for an interview is made, and final selection where admission is decided. Our study contributes to different streams of literature on selection procedures, i.e., labor economics, occupational psychology, and human resources, in a number of ways. First, we show that the latent criterion variables that we uncovered during non-participatory observation, interviews with jury members, and the RABK management team predict selection to the program. We are the first to identify such implicit selection criteria qualitatively and estimate their influence on the admission decision and future artistic performance using advanced econometric methods.

Second, we empirically show that the pre-selection round is useful in identifying inauspicious applicants. We demonstrate this by estimating the probability of an applicant being invited for an interview in relation to their succeeding performance. Our argument is that future performance is encapsulated in the applicants' potential the jury aims to identify, assuming success is independent of the program. We show that the 2016 artistic prestige ranking measured by Artfacts. Net improves as a consequence of pre-selection, going from 36,063 for all applicants to 28,164 for applicants invited for an interview (the same pattern is visible in the 7-year average rank). This improvement covers roughly 75% of the distance between deciles in the ranking distribution for all applicants. This is a significant contribution because the pre-selection round is considerably less costly in terms of resources devoted per applicant compared to final selection. At the same time, it brings a substantially larger gain in the average quality of the pool of candidates. We also show that the performance gain in artistic prestige rankings in the final selection round is statistically significant improving the average artistic prestige ranking by approximately 20% of the decile distance, going from 26,374 for the accepted applicants to 29,904 for the rejected ones (the same pattern is visible in the 7-year average rank).

Third, we simulate three alternative decision rules and compare them to the artistic prestige rankings of applicants accepted during final selection. We considered random assignment by lottery, weighted selection criteria, and biographical data. We found that actual selection outperforms the alternatives, although the gains are small. For instance, when we assess the benefits of intense jury deliberations during final selection and consider the realized allocation relative to random allocation or assignment by biographical data only, we show that the gains are minimal, i.e., less than 1%.

Finally, we estimate a switching regression model commonly used in labor economics, instead of a single linear equation for prediction of performance that is standard in other academic disciplines, such as occupational psychology and human resource management studies. We find a negligible selectivity bias emerging from unobservables in the selection equation that correlate with unobservables in the performance equations. This methodological consideration is important, especially concerning issues related to unobserved heterogeneity and selectivity bias in the estimated regression coefficients.

8 | CONCLUSION

In light of these contributions, our results directly question the benefits gained from rigorous and resource intensive selection procedures similar to those used in the final selection round of this study. That this costly round - based on sedulous assessments of applicants' future potential - does not result in striking differences in later performance among the accepted and the rejected applicants is in line with the related literature. We conclude that the difference between the two groups may be small for two reasons. First, this may be the case because the pre-selection round helps weed out the less promising applicants from a large pool of candidates, allowing only those judged to have exceptional potential to pass to the next selection hurdle. Second, the rejected applicants in the final selection round may apply – and be accepted - to other art programs[14] or find suitable alternatives, such as grants or stipends from different sources.

Finally, we empirically demonstrate a large gap in effectiveness between the first, more cursory, and cost-efficient pre-selection round versus the rigorous and financially intensive final selection. We propose two relevant considerations as to why other or less costly methods are not used. First, from the perspective of the selection or hiring committee, we maintain that paying attention to selection gives a sense of control under circumstances of uncertainty. To select by fast and simple methods, such as a lottery or simply tossing a coin, would suggest carelessness or indifference. Second, the effects of self-selection. In other words, a superficial selection procedure may tempt many applicants to take a chance. In contrast, a visibly costly one will activate more restraint and presumably lead to only higher-quality applicants applying (see Altangerel, 2019 for a formal model on self-selection under asymmetric information).

ACKNOWLEDGMENTS

The authors would like to thank Els van Odijk for access to the RABK data files and Martijntje Hallmann for generous help with collecting and interpreting the data. We would also like to thank the participants at the Thijs ten Raa Symposium, the ACEI Conference in Valladolid 2016, the EGOS Conference 2017 in Copenhagen and the Royal Netherlands Academy of Sciences (in particular from discussant Jan Willem Gunning) as well as Brigitte ten Brink, Edwin van Hooft and Paul Sackett for their constructive comments on earlier versions of this paper. We are grateful to the Kyklos Editors and anonymous reviewers for their insightful feedback during the review process.

Monika Kackovic would like to thank The Netherlands Organization for Scientific Research (NWO) for financial support from the Mozaïek Grant No. 017.008.111 and Joop Hartog would like to thank the Spanish Ministry of Education, Grant No. ECO2010-20829.

REFERENCES

Altangerel, K. (2019). *Essays on immigration policy*. CentER, Center for Economic Research.
ArtFacts.Net.net (2019). Homepage, (accessed April 20, 2019), [available at http://www.ArtFacts.Net.net].
Artnet.com (2019). Homepage, (accessed April 20, 2019), [available at http://www.artnet.com/].

[14]We randomly sampled 10% of the post-interview rejected candidates to better understand the efficacy of the focal final selection round, especially in relation to observed differences in success. We found that 92.5% were accepted to other post-graduate art programs.

Bille, T., Løyland, K., & Holm, A. (2017). Work for passion or money? Variations in artists' labor supply. *Kyklos*, *70*(3), 347–380. https://doi.org/10.1111/kykl.12141

Booth, A. L., & Satchell, S. E. (1995). The hazards of doing a PhD: an analysis of completion and withdrawal rates of British PhD students in the 1980s. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *158*, 297–318. https://doi.org/10.2307/2983293

Cascio, W. (1999). *Costing human resources*. South-Western College Publishing.

Castilla, E. J., & Rissing, B. A. (2019). Best in Class: The Returns on Application Endorsements in Higher Education. *Administrative Science Quarterly*, *64*(1), 230–270. https://doi.org/10.1177/0001839218759965

Caves, R. E. (2000). *Creative industries: Contracts between art and commerce*. Harvard University Press.

Crane, D. (2009). Reflections on the global art market: implications for the Sociology of Culture. *Sociedade e Estado*, *24*(2), 331–362. https://doi.org/10.1590/S0102-69922009000200002

Dana, J., Dawes, R., & Peterson, N. (2013). Belief in the unstructured interview: The persistence of an illusion. *Judgment and Decision making*, *8*(5), 512–520.

DiMaggio, P. (1987). Classification in art. *American Sociological Review*, *52*, 440–455. https://doi.org/10.2307/2095290

Ehrenberg, R., & Mavros, P. (1995). Do students' financial support patterns affect their times-to- degree and completion probabilities? *Journal of Human Resources*, *30*, 581–609. https://doi.org/10.2307/146036

Ertug, G., Yogev, T., Lee, Y. G., & Hedström, P. (2016). The art of representation: How audience-specific reputations affect success in the contemporary art field. *Academy of Management Journal*, *59*(1), 113–134. https://doi.org/10.5465/amj.2013.0621

Frey, B. (2019). The Artists' Labour Market. In *Economics of Art and Culture* (pp. 31–38). Springer.

Frey, B., & Pommerehne, W. (1989). *Muses and Markets: Explorations in the Economics of the Arts*. Blackwell.

Gallant, A. R., & Nychka, D. W. (1987). Semi-nonparametric maximum likelihood estimation. *Econometrica: Journal of the Econometric Society*, *55*, 363–390. https://doi.org/10.2307/1913241

Ginsburgh, V. A., & Van Ours, J. C. (2003). Expert opinion and compensation: Evidence from a musical competition. *American Economic Review*, *93*(1), 289–296.

Glejser, H., & Heyndels, B. (2001). Efficiency and inefficiency in the ranking in competitions: The case of the Queen Elisabeth Music Contest. *Journal of Cultural Economics*, *25*(2), 109–129. https://doi.org/10.1023/A:1007659804416

Grove, W., & Wu, S. (2007). The search for economics talent: doctoral completion and research productivity. *American Economic Review: Papers and Proceedings*, *97*, 506–511. https://doi.org/10.1257/aer.97.2.506

Hartog, J., & Kackovic, M. (2019). On the Idiosyncrasies of the Labour Market for Visual Artists: Striking features, a formal model, and suggestions for further work. *Labour*, *33*(2), 162–186. https://doi.org/10.1111/labr.12140

Heckman, J. (1974). Shadow prices, market wages, and labor supply. *Econometrica: Journal of the Econometric Society*, *42*, 679–694. https://doi.org/10.2307/1913937

Heckman, J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In *Annals of Economic and Social Measurement, Volume 5, number 4* (pp. 475–492). NBER.

Highhouse, S. (2008). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology*, *1*(3), 333–342. https://doi.org/10.1111/j.1754-9434.2008.00058.x

Kackovic, M., & Wijnberg, N. M. (2020). Artists Finding Galleries: Entrepreneurs Gaining Legitimacy in the Art Market. *Entrepreneurship Theory and Practice*, 1042258720950601. https://doi.org/10.1177/1042258720950601

Kausel, E. E., Culbertson, S. S., & Madrid, H. P. (2016). Overconfidence in personnel selection: When and why unstructured interview information can hurt hiring decisions. *Organizational Behavior and Human Decision Processes*, *137*, 27–44. https://doi.org/10.1016/j.obhdp.2016.07.005

Khaire, M. (2015). Art without borders? Online firms and the global art market. In O. Velthuis & B. Curioni (Eds.), *Cosmopolitan canvases* (pp. 102–128). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198717744.003.0005

Lievens, F., Highhouse, S., & De Corte, W. (2005). The importance of traits and abilities in supervisors' hirability decisions as a function of method of assessment. *Journal of Occupational and Organizational Psychology*, *78*(3), 453–470. https://doi.org/10.1348/096317905X26093

Marconi, G. (2018). Education as a long-term investment: The decisive role of age in the education-growth relationship. *Kyklos*, *71*(1), 132–161. https://doi.org/10.1111/kykl.12165

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, *79*(4), 599–616. https://doi.org/10.1037/0021-9010.79.4.599

Newey, W. (2009). Two-step series estimation of sample selection models. *The Econometrics Journal*, *12*, 217–229. https://doi.org/10.1111/j.1368-423X.2008.00263.x

Ones, D. S., Dilchert, S., Viswesvaran, C., & Salgado, J. F. (2010). Cognitive abilities. In J. L. Farr & N. T. Tippins (Eds.), *Handbook of employee selection* (pp. 255–275). Routledge/Taylor & Francis Group.

Prinz, A., Piening, J., & Ehrmann, T. (2015). The success of art galleries: A dynamic model with competition and information effects. *Journal of Cultural Economics*, *39*(2), 153–176. https://doi.org/10.1007/s10824-014-9217-2

Rijksakademie van beeldende kunsten (2019). Homepage, (accessed April 16, 2019), [available at http://www.RABK.nl]. Annual report (2017) available at: http://www.rijksakademie.nl/ENG/rijksakademie/profiel/jaarverslagen/

Rynes, S. L., Brown, K. G., & Colbert, A. E. (2002). Seven common misconceptions about human resource practices: Research findings versus practitioner beliefs. *Academy of Management Executive*, *16*(3), 92–103.

Schmitt, N. (2014). Personality and cognitive ability as predictors of effective performance at work. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*(1), 45–65. https://doi.org/10.1146/annurev-orgpsych-031413-091255

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Singer, L., & Lynch, G. (1994). Public choice in the tertiary art market. *Journal of Cultural Economics*, *18*, 199–216. https://doi.org/10.1007/BF01080226

Terpstra, D. E. (1996). The search for effective methods. *HR Focus*, *73*(5), 16–17.

Velthuis, O. (2013). *Talking prices*. Princeton University Press.

Velthuis, O., & Curioni, S. B. (2015). *Cosmopolitan canvases: the globalization of markets for contemporary art*. Oxford University Press.

Wiesner, W. H., & Cronshaw, S. F. (1988). A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview. *Journal of Occupational Psychology*, *61*(4), 275–290. https://doi.org/10.1111/j.2044-8325.1988.tb00467.x

Wright, P. M., Lichtenfels, P. A., & Pursell, E. D. (1989). The structured interview: Additional studies and a meta-analysis. *Journal of Occupational Psychology*, *62*(3), 191–199. https://doi.org/10.1111/j.2044-8325.1989.tb00491.x

Yogev, T., & Grund, T. (2012). Network dynamics and market structure: The case of art fairs. *Sociological Focus*, *45*(1), 23–40. https://doi.org/10.1080/00380237.2012.630846

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Kackovic, M., Hartog, J., van Ophem, H., & Wijnberg, N. (2022). The promise of potential: A study on the effectiveness of jury selection to a prestigious visual arts program. *Kyklos*, *75*(3), 410–435. https://doi.org/10.1111/kykl.12303