Can the Teacher Professional Education Admission Criteria in Indonesia Predict Teacher Performance?

by Asri Yusrina, Emilie Berkhout, Daniel Suryadarma, and Luhur Bima
RISE Indonesia

Introduction

Studies find that effective teachers raise student test achievement and lead to higher future earnings for the students (Chetty et al., 2014; Hanushek, 2011). Teacher selection and the criteria used in making the selection are important because they aim to identify such effective teachers. Identifying teachers with such potential is relatively more cost-effective than other policies applied after the teachers have teaching jobs (Klassen and Kim, 2019; Hobson et al., 2010). Many studies focus on selecting teachers based on the information collected at the time of hire to predict student outcomes (Jacob et al., 2018; Hill et al., 2012; Staiger and Rockoff, 2010). Other studies identify potentially effective teachers even before they become teachers. Those studies use information from teacher education programme admission criteria to predict teacher candidates’ success in the programme (Heinz, 2013; Casey and Child, 2011; Caskey et al., 2001). Among teacher selection criteria, studies identified predictors of subsequent performance including undergraduate grades, written tests, interviews, and teaching practice.

In developing countries, studies on teacher selection are virtually non-existent. We found two studies that focus on the selection of teachers during hiring. Both use candidates’ screening tests results to predict student learning outcomes (Araujo et al., 2020; Cruz-Aguayo et al., 2017). However, we did not find studies in developing country contexts that focus on selection of teachers into education programmes or how the admission criteria relate to student learning outcomes.

Key Points

• A rigorous teacher selection process is important because it leads to a better teacher workforce. For developing countries with more limited resources, recruiting effective teachers could be more cost-effective compared to other policies to improve learning outcomes.

• The admission criteria of Indonesia’s flagship pre-service teacher training programme, the Teacher Professional Education or PPG, include having an undergraduate GPA above 3; taking the computer-based online admission tests, which include an aptitude test, a pedagogical knowledge test, an English proficiency test; and attending an interview.

• The admission criteria are significant predictors of teacher candidates’ success at the end of their professional education.

• The same admission criteria did not predict teacher performance in a real classroom setting as measured by student learning outcomes.

• We conclude that predicting student learning outcomes based on the admission criteria is unreliable due to the lack of pre-test scores of students. Further research should explore whether the tests administered in the programme can be improved to better identify teachers that will impact student learning outcomes or if a different selection method can be used to select teachers.
Whether focusing on selecting teachers during their education programme or as they go through the recruitment process, studies on teacher selection across countries have the same underlying question: Will the criteria be able to identify effective teachers? The idea of teacher selection to improve the quality of the teaching force is appealing. For instance, in high performing countries in PISA, like Japan and Korea, where there are many teacher colleges (Ingersoll, 2007) and the most prevalent teacher employment is civil-service, great attention is paid to the quality of selection into teacher education programmes (OECD, 2018).

Teacher selection is arguably more critical in developing countries. In most developing countries, the entry into teacher education programmes lacks selectivity and teacher qualifications tend to be set lower compared to other professional jobs (Béteille and Evans, 2019). Across all developing countries, a larger number of teachers are employed and account for most of the education spending, but their effect on student outcomes is small (ADB, 2021; Crawfurd and Pugatch, 2021). This suggests the need for more attention to policies such as the selection of teachers and criteria used to identify those best suited to teach in the classroom.

In Indonesia, where the teacher recruitment system lacks a strong mechanism to ensure quality (Huang et al., 2020) and the teacher in-service training has not been effective (Revina et al., 2020), a potential way to improve the pool of teachers is through enhanced selection of individuals who will become teachers. We specifically question whether we can predict a teacher’s performance using information available when they were a teacher candidate. Admission criteria for teacher education are presumably intended to identify candidates who have the greatest likelihood of being able to do well in the academic programme and ultimately in the classroom as a professional. The identification of criteria that predict teacher subsequent performance would give policy makers a stronger understanding of where programme improvement may be needed.

Overview of the PPG Programme

In this Insight Note, we use data from a highly selective government-run pre-service teacher training programme in Indonesia, Pendidikan Profesi Guru (PPG), or Teacher Professional Education, to examine the correlation between performance in the selection process and performance in the summative assessment. We also examine the correlation between the former and the learning outcomes of students that PPG graduates teach.

The Pre-Service PPG is an additional one-year programme to the four-year bachelor’s degree in education (see Figure 1). After PPG graduation, teacher candidates are not guaranteed jobs. They have employment options to become a civil servant teacher, a private school teacher, or an honorary contract teacher.

Figure 1. Pathways to Teacher Professional Education Programme and Teaching
The Pre-Service PPG was launched because the government was concerned that most teacher colleges were producing subpar graduates (Direktorat Jenderal Pembelajaran dan Kemahasiswaan, 2018). Subsequently, the PPG was created to ensure minimum teacher quality standards were met. The programme was rolled out as a pilot in 2017—out of 421 teacher colleges in Indonesia, only 23 were allowed to offer the PPG programme that year and 29 teacher colleges in the next year. The government intends to use PPG as a teacher licensing programme in the long-term.

The PPG programme attracts undergraduate students because it offers the following benefits. First, PPG graduates will automatically receive a perfect score on the competency assessment of teacher civil-service recruitment. Thus, the likelihood of being hired as a civil service teacher is greater for PPG graduates compared to other applicants. Since most teachers aspire to become civil servants (Jalal et al., 2009), such an advantage makes the PPG very attractive. Second, PPG graduates will automatically pass the government’s teacher certification phase. Although not a teaching requirement, the teacher certification is sought because it provides teachers with an allowance that essentially doubles their take-home pay (De Ree et al., 2018).

The selection of teacher candidates into the PPG programme is carried out through successive filters. The first is the standard administrative requirements that candidates must meet, such as having an age of 31 years or less, an undergraduate GPA above 3 (on a scale of 0–4) from an accredited university, and less than five years of teaching experience before attending PPG. The second filter is the computer-based online admission tests, which include an aptitude test, a pedagogical knowledge test, and an English proficiency test. Candidates who pass all the online admission tests—measured by the weighted average of the three scores—are invited for an interview with the selectors. The interview and the overall online admission test scores are weighted at 30 percent and 70 percent, respectively, to obtain the final score. Candidates with a final score above 60 are admitted into the Pre-Service PPG programme. At the end of the programme, they must pass two exit exams covering knowledge and teaching practice to graduate.

Figure 2 shows the admission selection process of the PPG programme for primary school teachers in twenty-nine teacher colleges in 2018. The proportion of teacher candidates admitted into the Programme over applicants who scored above 50 in the online admission test was about 87.2 percent. Candidates admitted into PPG had to re-register to confirm their willingness to enrol in the Programme. We found that the number of graduates was less than the number of candidates enrolled. However, it was not because they failed the exams but because some candidates opted out of the Programme. On average, one candidate in each teacher college opted out of the PPG.

Of the teacher candidate cohort that we study, who enrolled in 2018, the passing rate of the exit exam on the first attempt is 94.27 percent of 1291 candidates. Teacher candidates who failed the exam have another chance to repeat the exam.

Figure 2. The PPG Programme Admission Selection Process

---

1 In December 2020, we carried out a phone survey of teacher candidates who applied to the PPG programme (admitted and not admitted). About 45 percent of the 314 candidates we contacted said they had worked as civil servant teachers.
Does Performance during the PPG Admission Phase Predict Performance at the End of the Programme?

We analyse the administrative dataset on the universe of 1,291 teacher candidates admitted into the PPG for primary school teachers in Indonesia. In Table 1, we find that the candidate's performance on some admission tests is a significant predictor of their performance in the exit exams. When controlling for other factors, a one standard deviation higher combined score on the online admission test is associated with a 0.3 standard deviation higher score on the knowledge exam. A one standard deviation higher score on the interview is associated with a 0.07 standard deviation higher score on the teaching practice exam.

We also find that the undergraduate GPA is a significant predictor of performance on both exit exams compared to the interview score. A one standard deviation higher undergraduate GPA, on average, is associated with a 0.17 standard deviation higher score on the knowledge exam and a 0.06 standard deviation higher score on the practice exam.

Table 1. Regression Results between Admission Criteria and Exit Examination

<table>
<thead>
<tr>
<th></th>
<th>(1) Standardised Score on Knowledge Exam</th>
<th>(2) Standardised Score on Teaching Practice Exam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standardised Overall Online Admission Test Score</td>
<td>0.30*** (0.03)</td>
<td>0.04 (0.03)</td>
</tr>
<tr>
<td>Standardised Interview Score</td>
<td>0.03 (0.03)</td>
<td>0.07*** (0.03)</td>
</tr>
<tr>
<td>Standardized Undergraduate GPA</td>
<td>0.17*** (0.03)</td>
<td>0.06*** (0.02)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.42** (0.21)</td>
<td>0.43*** (0.15)</td>
</tr>
<tr>
<td>Teacher colleges dummy</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1291</td>
<td>1291</td>
</tr>
<tr>
<td>R²</td>
<td>0.30</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Does Performance during the PPG Admission Phase Predict Student Learning Outcomes?

Our subsequent analysis sought to examine the relationship between performance during the PPG admission phase, exit phase, and student learning outcomes. We were able to reach, via phone survey, 121 PPG graduates from seven teacher colleges. We could only visit 113 homeroom teachers since there are teachers who move to different schools and cannot be reached through the phone, and some schools denied the data collection. Since we carried out the data collection during Covid-19 pandemic, schools were not having face-to-face meetings. The schools’ closure hindered us in conducting tests of all students taught by the sample teachers. We sampled 15 students of the total students taught by the 113 teachers. We administered the student tests by visiting individual students’ houses with a strict health protocol in the data collection. In total there are 1,560 students. When we collected data on student learning outcomes from February to April 2021, the sample teachers had been teaching for six months in the 2019/2020 academic year, approximately two years after graduating from the PPG.
Our results show that the teachers’ performance in the PPG admission tests is not correlated with their students’ learning outcomes. We also use candidates’ exit exams score to predict student learning outcome and the results are not significant. The insignificant results should be taken cautiously since we do not have pre-test scores of the students that control students’ prior knowledge.

Table 2: Regression Results between Admission Criteria, Exit Examinations and Student Learning Outcome

<table>
<thead>
<tr>
<th></th>
<th>(1) Standardised Score on Numeracy</th>
<th>(2) Standardised Score on Numeracy</th>
<th>(3) Standardised Score on Literacy</th>
<th>(4) Standardised Score on Literacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standardised Overall Online Admission Test Score</td>
<td>-0.024 (0.033)</td>
<td></td>
<td>-0.003 (0.016)</td>
<td></td>
</tr>
<tr>
<td>Standardised Interview Score</td>
<td>0.015 (0.036)</td>
<td></td>
<td>0.038 (0.029)</td>
<td></td>
</tr>
<tr>
<td>Standardized Undergraduate GPA</td>
<td>0.002 (0.006)</td>
<td></td>
<td>0.002 (0.005)</td>
<td></td>
</tr>
<tr>
<td>Standardised Score on Knowledge Exam</td>
<td></td>
<td>0.007 (0.007)</td>
<td></td>
<td>0.005 (0.004)</td>
</tr>
<tr>
<td>Standardised Score on Teaching Practice Exam</td>
<td></td>
<td>-0.003 (0.004)</td>
<td></td>
<td>-0.004 (0.004)</td>
</tr>
<tr>
<td>Student's gender (1=female)</td>
<td>0.108*** (0.039)</td>
<td>0.110*** (0.040)</td>
<td>0.193*** (0.038)</td>
<td>0.195*** (0.038)</td>
</tr>
<tr>
<td>Student takes a private lesson</td>
<td>0.083 (0.070)</td>
<td>0.084 (0.070)</td>
<td>0.057 (0.052)</td>
<td>0.058 (0.052)</td>
</tr>
<tr>
<td>Student’s housing quality index</td>
<td>0.032* (0.012)</td>
<td>0.032* (0.012)</td>
<td>0.028 (0.016)</td>
<td>0.030 (0.016)</td>
</tr>
<tr>
<td>School quality index</td>
<td>-0.000</td>
<td>0.000</td>
<td>-0.001</td>
<td>-0.001</td>
</tr>
<tr>
<td>Classroom average score in Literacy or Numeracy</td>
<td>0.979*** (0.012)</td>
<td>0.980*** (0.012)</td>
<td>0.992*** (0.013)</td>
<td>0.991*** (0.013)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.086</td>
<td>-0.085***</td>
<td>-0.169***</td>
<td>-0.121***</td>
</tr>
<tr>
<td>Grade dummy</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1545</td>
<td>1530</td>
<td>1543</td>
<td>1528</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.513</td>
<td>0.512</td>
<td>0.368</td>
<td>0.369</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.509</td>
<td>0.508</td>
<td>0.362</td>
<td>0.364</td>
</tr>
</tbody>
</table>

Notes: The number of observations falls from 1,560 because missing data on the school quality index or student characteristics. Standard errors in parentheses and clustered at the teacher level. * p < 0.05, ** p < 0.01, *** p < 0.001

Prior to data collection of the 113 teachers, we attempted to collect data of the targeted sample when they were homeroom teachers in the 2019/2020 school year. Due to the COVID-19 surge and lockdown measures we could not complete our data collection of the targeted sample. We were able to visit 32 teachers of PPG graduates and carried
out live observations when they were teaching in the classroom. The observers were to record activities that took place in the classroom every five minutes. Depending on the length of the lesson, each teacher had different total observation in the range of 7 to 20 data points. To combine the data from each class, the actual time was not used but rather the relative time from 0 to 100 percent of class time. For example, for the class of 40 minutes, equal to 8 data points, if the teacher’s presentation in front of the class occurred for 10 minutes (2 data points) then 25 percent of the time was used for whole-class interaction.

We use the classroom observation data to seek further if the unreliability of predicting student learning outcome derives from the fact that there are no differences in the teaching practice of the teachers. The objective of the analysis is to examine teaching practices across the admission criteria. We want to know if there is a difference in the classroom setting or teaching practices of the highest 20 percent and the lowest 20 percent in the online admission score and interview score.

Figure 3: Types of Classroom Setting based on Teacher’s Online Admission Score

![Figure 3: Types of Classroom Setting based on Teacher’s Online Admission Score](image)

Figure 4. Types of Classroom Setting based on Teacher’s Interview Score

![Figure 4. Types of Classroom Setting based on Teacher’s Interview Score](image)

We cannot identify different patterns of teachers’ classroom setting based on the selection criteria. However, teachers in the highest 20 percent of online admission test and interview score dedicated more time for the whole-class setting. There is no significant difference in terms of proportion of time for each type of classroom setting between teachers in the highest 20 percent and lowest 20 percent of the online admission test score or interview test score. Strikingly, the transition between classroom setting occupies about 20 percent of the lesson across all sample. The transition means a switch from whole-class to groupwork or individual work and vice versa that occurs within five minutes of the observation record.

---

2 There are 23 of 32 teachers who were revisited in the 2021 data collection or part of the 113 teachers.
To gauge teacher’s teaching practices that boost students’ understanding (Bruns and Luque, 2014; Siraj and Taggart, 2014; Lee and Kinzie, 2011), we analyse the use of learning materials, the occurrence of teacher’s effort in making connection between lessons, and asking open-ended questions.

Figure 5. Teaching Practices based on Teacher’s Online Admission Score

Figure 6. Teaching Practices based on Teacher’s Interview Score

Are the Admission Criteria Appropriate for Selecting Teachers?

We conclude that the admission criteria are significant predictors of a teacher candidate’s success at the end of the PPG programme. The overall online admission test score indeed predicts a candidate’s success in the knowledge exam, whereas the interview only predicts their success in the teaching practice exam. This finding is similar to Heinz (2013), who found that a selector’s judgment in an interview correlates with a candidate’s teaching practice performance. Heinz argued that the interview typically looks for personality traits representing teaching abilities valued in the practice exam rather than academic abilities.

One might assume that there is lower reliability in the exit exams since all candidates eventually graduate. While we did not have information about the validity of the knowledge exam, the teaching practice exam was designed to avoid

---

5 Using the term “citizen” here is not intended to suggest that the views and needs of non-citizen residents are not important in the politics relationship. Instead, this is partly a term of convenience and partly a term to highlight the importance of electoral politics in many contexts.
favouritism. The evaluators involved in scoring of the teaching practice exam were not the candidate’s supervisor or mentor.

Although undergraduate GPA predicts both exit exams, it is less a predictor of candidate’s performance at the end of the programme than the online admission test and interviews. This is likely due to the lack of uniformity in grading practices between teacher colleges. Similar to Caskey et al (2001), GPA shows the lowest correlation with measures of candidate’s performance in the end of the programme the compared to other admission criteria. We also found that the correlation between undergraduate GPA with the teaching practice exam is smaller than that with the knowledge exam. However, it also shows that the undergraduate GPA is just as good in predicting teaching practice.

While we can answer that admission criteria are associated with success in the PPG programme, we find no correlation between those measures and student learning outcomes. Based on this finding, we conclude that predicting student learning outcomes using admission criteria is unreliable due to the lack of pre-test scores of students.

Potentially, the insignificant correlation may also come from the fact that there are no differences in the teaching practices of the teachers. Using classroom observation data of 32 sample teachers who are PPG graduates, we found no differences in teaching practices across distribution of online admission test scores and interview scores. One striking piece of evidence shows that the time used to transition between classroom settings (i.e., from whole-class to groupwork or individual work and vice versa) occupies about 20 percent of the lesson time across all samples. More time needed to do the switch indicates inefficient transitions and a loss of time that could be dedicated to instructional time.

Although there are no differences in teaching practices across distribution of admission criteria, it is possible that sample teachers of PPG graduates have yet to perform a well-defined way of teaching at the time of the observation. And whether the classroom practices have any influence on student performance is beyond our analysis.

The next question remains as to whether the tests administered in the programme can be improved to better identify teachers that will impact student learning outcomes, or if a different selection method can be used to select teachers.

One improvement that can be considered is to include a subject content knowledge test, particularly in numeracy and literacy content, in the admission tests of the PPG programme. Subject content knowledge might serve as a more stable predictor of a candidate’s success (Caskey et al., 2001) and a better predictor of students’ learning outcomes (Ragatz, 2015).

Pedagogical knowledge tests may also be used in the admission tests of the PPG programme. The evidence from the Indonesian TIMSS video study suggests that pedagogical knowledge tests predict student learning outcomes, although not as well as a subject content knowledge test (Ragatz, 2015). The TIMSS study noted that translating the pedagogical concepts into a written test presents difficulties. As a consequence, there is a chance that a pedagogical knowledge test is unable to gauge a teacher’s performance when it is administered in the selection process.

In addition to a set of measures used to screen teacher candidates, Staiger and Rockoff (2010) suggest that a probationary period within the first few years of teaching is likely to be a better strategy when selecting good teachers. More broadly, this study highlights the need for further research on the teacher selection criteria and strategy that best predict teacher performance in real classroom settings.
References


Ingersoll, R. 2007. A Comparative Study of Teacher Preparation and Qualifications in Six Nations. CPRE Research Reports. [online] Available at: https://repository.upenn.edu/cpre_researchreports/47


Asri Yusrina is a senior researcher on the RISE Indonesia team. She completed her master’s degree in Economics at the University of Queensland, Australia. Setting her focus on applied econometrics in microeconomics, she wrote a thesis about the effect of teacher certification program on the absence status of civil service teachers. Her knowledge in education research has been deepened through her involvement in the Study on Teacher Absenteeism and Independent Impact Evaluation of the KINERJA Program. She has experiences in managing surveys and is knowledgeable on Indonesian datasets such as National Socio-Economic Survey (Susenas).

Emilie Berkhout is an international consultant for the RISE Indonesia team based at the Amsterdam Institute for Global Health and Development (AIGHD). She obtained her Master’s Degree in Development Economics with distinction at the VU University in 2015. She conducts impact evaluations to assess effects of policy interventions in developing countries. Her main research focuses on the effect of policy reforms in basic education on learning outcomes in Indonesia.

Daniel Suryadarma is a research economist at the Asian Development Bank Institute in Tokyo. He is an applied microeconomist, conducting policy experiments and analyses in the areas of education, poverty, and social policy with government and non-government partners. My work has appeared in high profile peer-reviewed economics journals, including American Economic Journal: Applied Economics and Journal of Development Economics. My current research spans South Asia, Southeast Asia, and the Pacific.

Luhur Bima is a senior researcher on the RISE Indonesia team. He holds a master’s degree in economics from Uppsala University in Sweden and a bachelor’s degree in international economics and business from the University of Groningen in the Netherlands. He worked with Bappenas and completed an internship with the University Medical Centre in Groningen, the Netherlands. Bima had involved in various research projects, including A Study on Teacher Absenteeism, Independent Impact Evaluation of the KINERJA Program, Child Poverty and Disparities in Urban Area, and Multidimensional Child Poverty.

Citation: