

Pre-Analysis Plan: Humans versus Smartbots: Scaling-up behavioral interventions to reduce teacher shortages *

Nicolás Ajzenman [†] Gregory Elacqua [‡] Catalina Hermosilla
Graciela Pérez-Núñez [§] Ana Teresa del Toro Mijares [¶]

November 2024

Abstract

We design an experiment to assess the relative effectiveness of human mentors versus AI-powered chatbots in motivating students to pursue a teaching career in Chile. We work with Elige Educar, a Chilean NGO that implements a mentoring program for students interested in education or related fields. Participating high school students are randomly assigned to one of three groups: a group that receives support from a human mentor via WhatsApp; a group that receives support from an AI-powered chatbot, also through WhatsApp; and a passive control group that receives no support. We assess the effectiveness of each treatment in motivating students to register for an education major through Chile's centralized university admissions process. Beyond addressing Chile's immediate teacher recruitment challenges, the study contributes to the emerging research on AI in education by empirically assessing the potential of generative AI to provide personalized, scalable mentoring.

JEL classification: D91, I23, I25

Keywords: Teachers, teacher policy, teacher shortages, scale-up, behavioral, chatbots

*We thank the Elige Educar team for their collaboration in this study. This research was approved by the HML Institutional Review Board (Study #2713).

[†]McGill and IZA. E-mail: nicolas.ajzenman@mcgill.ca

[‡]Inter-American Development Bank. E-mail: gregorye@iadb.org.

[§]International Institute for Education Planning, UNESCO. E-mail: gnperez@gmail.com.

[¶]Inter-American Development Bank. E-mail: anade@iadb.org

1 Introduction

1.1 Context and motivation

Chile currently faces a severe shortage of qualified teachers, driven by an 8% annual decrease in teacher education program enrollment over the last five years coupled with high teacher attrition. By 2025, Chile is projected to have a deficit of over 30,000 teachers, particularly in math and science, and in disadvantaged regions and schools ([Medeiros et al., 2018](#); [Elige Educar, 2020](#)).

Chile has implemented various measures to address teacher shortages and improve the quality of the teaching profession. The *Reforma Educacional* enacted in 2016, the centerpiece of these efforts, introduces a merit-based teacher career path with monetary incentives for high-performing teachers, especially those who work in disadvantaged contexts. Policies have also been introduced to attract high-performing students to the teaching profession by raising admission standards for education programs and increasing accreditation requirements for these programs and the institutions that offer them. Additionally, the government has launched initiatives to support teachers in their early years of professional practice, including through the establishment of a National Induction and Mentoring System in 2017. Other interventions have been developed by non-state actors such as the NGO Elige Educar, which has developed a program “Quiero Ser Profe” with the aim of increasing teacher recruitment, by providing support and information to high school students who are considering a career in education.

Despite these ongoing efforts, challenges in teacher recruitment and retention persist. Many of the most significant policies –such as merit-based pay increases, teacher induction and mentoring programs, advocacy campaigns and support initiatives like the “Quiero Ser Profe” program– have been impactful, but have had limited reach because they are costly or require highly specialized human resources to scale. Finding ways to increase the scale of these programs while maintaining effectiveness is critical to addressing the teacher shortages prevalent in Chile and elsewhere in the region.

To this end, the IDB partnered with Elige Educar in 2020 in a project that evaluated the effectiveness of human mentors trained by Elige Educar in supporting high school stu-

dents considering an education career through the “Quiero Ser Profe” program, compared to a rule-based chatbot designed to provide similar support. While the pilot demonstrated that chatbot technology could effectively reduce costs of labor-intensive initiatives and increase scalability of such programs, both treatment arms showed positive results in motivating high school students to apply for teacher education programs. However, the effect of the rule-based chatbot was limited and not statistically significant, mainly due to the limited engagement and interaction possible with the user through the rule-based chatbot technology available at the time ([Ajzenman et al., 2024](#)).

Today the emergence of generative AI technologies offers a significant opportunity to enhance the reach and impact of programs such as “Quiero Ser Profe,” without compromising their effectiveness. To this end, this study builds on the prior work by complementing the 2020 study, now using an AI-powered chatbot.

1.2 Study objectives

Our primary research question is: *To what extent are AI-powered chatbots able to emulate the support provided by trained human mentors, in mentoring and motivating high school students to consider applying for teacher education programs in Chile?* As secondary research questions, this study also aims to understand to what extent this effectiveness is mediated by the intensity, duration and type of interactions between the student and the mentor or AI-powered chatbot, and to what extent this effectiveness differs according to student characteristics including gender and achievement.

The experiment consists of providing personalized mentoring through WhatsApp to students who are applying to Chile’s centralized university admission process and have expressed interest in education or related careers. This mentoring support seeks to resolve their doubts and motivate them to enroll in education programs. The experiment randomly assigns high school seniors to one of three arms, drawing from two distinct groups: those who explicitly express interest in education careers and those interested in careers in the social sciences but not specifically in education. The three experimental arms are: (1) human mentoring, where mentors (mostly college-level education students) trained in the “Quiero Ser Profe” methodology contact students via WhatsApp to inform them about education programs, scholarships, and teaching career benefits; (2)

AI-powered chatbot mentoring, where chatbots trained to provide similar support to the human mentors perform the same task; or (3) a passive control group that receives no contact. Both human and chatbot mentors follow specific guidelines and provide only factual and truthful information, with no deception involved.

Our primary outcome of interest is the probability of students applying to education majors, measured through administrative data from the university application process. Additionally, we will analyze students' interactions via WhatsApp with human mentors and with the AI-powered chatbots, to better understand the mechanisms behind any differences in effectiveness between the two treatment arms. By including both students already interested in education and those interested in social sciences more broadly, we can assess whether these interventions are effective in reinforcing existing preferences, in attracting new candidates to the teaching profession, or in both.

The novelty of this study lies in exploring the potential of AI-powered chatbots to make the "human" intervention scalable, addressing the current limitation: the lack of capacity and resources to deploy enough trained human mentors to scale the intervention across the entire relevant population. While a growing body of literature demonstrates that chatbots are not necessarily as persuasive as humans - despite often providing similar information - the effectiveness of chatbots in appearing credible and trustworthy remains an empirical question with mixed findings across different contexts. More broadly, the study contributes to the growing research on the effectiveness of AI-powered technologies in education, especially of their use and efficacy in scaling labor-intensive programs or policies. This project aims to examine their effectiveness as a low-cost solution to expanding the scale of programs focused on improving teacher recruitment such as "Quiero Ser Profe," given their added benefit of more closely emulating human interaction. We hypothesize this will result in improved engagement with users, ultimately leading to increased enrollment in teacher education programs.

The rest of this document proceeds as follows. In Section 2, we describe the intervention and the experiment's design. In Section 3, we present the relevant hypotheses and data. Section 4 concludes with a discussion of the empirical methodology that will be used to identify and estimate causal effects in this setting.

2 Experimental Design

2.1 The "Quiero Ser Profe" mentoring program

We work in collaboration with the Chilean NGO Elige Educar, which leads the implementation of the intervention. Elige Educar is an organization under the umbrella of the Pontificia Universidad Católica de Chile and has more than 15 years of experience in generating studies, interventions and projects associated with teachers and the teaching profession in Chile. Elige Educar's mission is to ensure all children have quality teachers, and the organization's work focuses on three strategic lines of action: (i) attraction and retention of high-performing students to teacher education programs; (ii) improvement of the professional and social valuation of teachers; and (iii) the promotion of public policies that improve teaching and teaching conditions. Most Elige Educar programs and initiatives are conducted in partnership with the Chilean Ministry of Education.

Since 2010, Elige Educar has implemented the program "Quiero Ser Profe" ("I Want To Be a Teacher" in English) with the aim of increasing recruitment into teaching careers in Chile. The program is geared towards Chilean students in their final year of high school. Participants for the program are recruited via the Department of Evaluation, Measurement and Educational Registry (DEMRE, by its Spanish acronym) which oversees the development, implementation and management of the centralized admissions process for higher education institutions. Elige Educar has a research agreement with DEMRE whereby the department shares the full database of all high school students who register for admission to the higher education system in each cycle, who have indicated that they are interested in education or in related careers, who consent for their data to be shared with Elige Educar, and who also consent to be contacted to participate in studies and research related to education. These students comprise the potential participants for the "Quiero Ser Profe" program.

A subset of these students are contacted via WhatsApp by an Elige Educar mentor who has been trained in the "Quiero Ser Profe" methodology to provide high-quality mentoring support, focused on motivating them to consider a career in education. Mentors are typically university students majoring in education who plan to become teachers. In

addition to individual mentoring, Elige Educar delivers a massive media campaign every year that reinforces the goal of attracting students to education majors. This media campaign, funded partially by the Chilean Ministry of Education, provides information about education careers using mass media outlets (television, radio, billboards on public roads) and social networks.

Amongst other topics, mentors receive online training to provide information and guidance on:

- Available education programs at the university level and their characteristics;
- Financing and scholarship options;
- Requirements for entrance to specific education programs, including information on necessary exams and scoring minimums;
- Work conditions and job market for education graduates; and
- Broader vocational guidance on the advantages of a career in education.

Mentors reach out to participating students via WhatsApp from late October through mid-December, and then again the first week of January prior to the university registration period closing on January 9.

The program has traditionally provided individual mentoring and information during the college application period exclusively to students who explicitly express interest in education careers. As an innovation in this version of the study, we expand the target population to include students interested in social sciences but not specifically in education. This expansion allows us to evaluate whether the program can effectively attract students who haven't initially considered teaching as a career path. While this represents the first time human mentors will interact with students not explicitly interested in education, the program's content and structure remain unchanged, maintaining its original design and implementation strategy.

2.2 Developing the "Quiero Ser Profe" smartbot

In collaboration with Elige Educar, we develop an Artificial Intelligence (AI)-powered chatbot designed to emulate the support and guidance provided by trained mentors.

We train the model using ChatGPT-4mini and a comprehensive database of resources approved by Elige Educar. This state-of-the-art language model is chosen for its ability to engage in natural conversations while maintaining tight control over the accuracy and appropriateness of responses.

The model is trained on a content bank of resources and materials that are used to train and prepare the Elige Educar mentors. The knowledge base provides authoritative information spanning the entire higher education application process, from university admission test details to specific program information. The database includes complete information about the *Prueba de Acceso a la Educación Superior* (Test to Access Higher Education, or PAES, by its Spanish acronym), Chile's standardized university entrance exam, including exam dates, topics covered, and requirements, as well as detailed data about teacher education programs offered by every university accredited by the National Accreditation Commission. The approved resources also cover financial aspects of pursuing a teaching degree, including comprehensive information about the *Beca Vocación de Profesor* (Teaching Vocation Scholarship) and other available funding opportunities. For each accredited program, the database contains specific details about tuition fees, program duration, and application requirements, ensuring prospective students have access to accurate and up-to-date information about their options. The model is trained to restrict the information provided to its knowledge base and it is not able to search the Web for additional information.

The AI-powered chatbot is developed and tested over a period of 12 weeks prior to its deployment in the experiment. It undergoes extensive technical testing to ensure that it is drawing accurate and relevant information from its knowledge base and that its guidance and advice is provided in an engaging and accessible manner to the target audience. Based on feedback from users in the initial testing and piloting phase, the chatbot undergoes additional development prior to deployment.

The chatbot is developed as "Kai," a non-binary character that is already part of Elige Educar's communication campaigns. This design choice was deliberate, building on Elige Educar's existing brand identity and communication strategy.

2.3 Intervention design

The randomized controlled trial (RCT) evaluates the relative effectiveness of two support delivery methods: human mentors and an AI-powered chatbot ("Kai").

Both treatment arms operate from Elige Educar's WhatsApp Business account, to help ensure that participants receiving the message trust its source. Initial message templates are developed for both treatment arms to ensure maximum similarity in the first contact. Human mentors introduce themselves by name and specify which teaching program they are studying. The AI chatbot introduces itself as "Kai," and identifies itself as Elige Educar's AI-powered chatbot. The initial message in both arms includes four interactive buttons: two provide additional information about the mentoring program, one initiates the mentoring session, and one allows participants to opt out of the program and be removed from the database. The complete messages used in each arm can be found in [Appendix A](#).

Following the initial message, the chatbot interaction proceeds freely based on student interests and needs. Human mentors follow a non-structured script that covers key topics needed for a successful conversation: they share more about themselves and their studies, get to know the student (including their background and academic status), explore motivations for considering teaching as a career, and provide motivational reinforcement linked to the student's specific interests through arguments and personal experiences. The mentors also address concerns, provide information about teaching programs, application requirements, program selection recommendations, and career financing options. Additionally, mentors ensure students know they can provide information about PAES exam dates and the application process.

Both arms follow a similar contact protocol: up to two attempts during the student's assigned intervention week. The first attempt introduces the program and asks about the student's interest in receiving guidance. If there is no response, a second and fi-

nal attempt is made to establish contact. The approach is customized based on student responses, informed by previous research showing that emphasizing different motivations can lead to unexpected outcomes if not tailored to the appropriate audience ([Ajzenman et al., 2021](#)). This streamlined contact protocol represents an optimization from a previous version of the experiment that used three attempts ([Ajzenman et al., 2024](#)).

In both treatment arms, participants who maintained any type of interaction receive a short survey asking for their satisfaction with the mentoring provided and requesting feedback. Additionally, participants who wish to proactively contact the mentor again with other queries or questions are able to do so.

The control group is passive and does not receive any support or messaging.

2.4 Timeline

The intervention spans from October 2024 to January 2025, strategically covering the period when students make their college application decisions.

Following student registration and consent collection through DEMRE in July 2024, the intervention begins the final week of October with weekly activation of student subgroups. Over the next eight weeks and through December 20, all students in both treatment arms receive their initial WhatsApp contact. Each student's regular intervention period includes two contact attempts during their assigned week, with interactions continuously logged for both arms.

A particularly intensive phase of support occurs during the week of January 6th, 2025, when students receive their PAES results and must make their university applications. During this critical period, both human mentors and the chatbot contact all students who haven't explicitly opted out of the program, regardless of their previous interaction levels. This special support phase includes specific guidance on the application process and step-by-step instructions for submitting applications effectively, ensuring students understand their options within education programs.

Following the application period, the formal intervention concludes in January 2025. Administrative data on applications and enrollment will be collected in March 2025,

with data analysis beginning in April after proper anonymization by DEMRE.

2.5 Sample Selection and Treatment Assignment

Our sample consists of students registering for the *Prueba de Acceso a la Educación Superior* (PAES), Chile’s standardized university entrance exam. During registration with DEMRE, students complete a survey indicating their academic interests. Within this survey, students can agree to be contacted by Elige Educar and authorize the use of their data in Elige Educar’s research initiatives. Of the 294,020 students registered for the PAES, 151,253 (51.4%) agreed to both conditions. From this pool, 23,690 students expressed interest in education careers, while 19,849 indicated interest in social sciences but not in education, resulting in a potential sample of 43,539 students.

The stratification process considers three key dimensions that previous research has identified as relevant to career choice: gender (male or female), educational background (based on type of high school: public, private-subsidized, or private non-subsidized), and initial career interests (education or social sciences). The combination of these variables generates twelve strata. Operational capacity constraints determine the final sample distribution:. The human mentoring arm is limited to 9,000 students, while the chat-bot arm can accommodate 20,000 students. The remaining 14,539 students are assigned to the control group. This allocation strategy ensures we maximize the use of available mentoring resources while maintaining a substantial control group for comparison.

Table 1: Baseline covariates summary statistics

	Obs	Mean	Std. Dev.	Min	Max
Female	43,539	0.650	0.477	0	1
<i>High school type</i>					
Public	43,057	0.361	0.480	0	1
Private-Subsidized	43,057	0.570	0.495	0	1
Private	43,057	0.070	0.254	0	1
<i>Interest</i>					
Education	43,539	0.544	0.498	0	1
Social sciences	43,539	0.456	0.498	0	1

Table 2: Covariates balance check

	Human mean	Bot mean	Control mean	Human vs Control diff	Bot vs Control diff	Human vs Bot diff
Female	0.649 (0.477)	0.650 (0.477)	0.649 (0.477)	0.000	0.001	-0.001
<i>High school type</i>						
Public	0.350 (0.479)	0.357 (0.479)	0.356 (0.479)	-0.006	-0.001	.007
Private-subsidized	0.569 (0.495)	0.569 (0.495)	0.571 (0.495)	-0.002	-0.002	0.000
Private	0.068 (0.252)	0.069 (0.254)	0.070 (0.256)	-0.002	-0.001	-0.001
<i>Interest</i>						
Education	0.552 (0.450)	0.544 (0.498)	0.540 (0.498)	0.012*	0.004	0.008
Social sciences	0.448 (0.497)	0.456 (0.498)	0.460 (0.498)	-0.012*	-0.004	-0.008

Notes: Standard deviation in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The sample is composed of 65 percent women, with 57 percent of students attending private-subsidized schools, which are owned and operated privately but receive per-student public subsidies and cater mainly to middle-class families. Approximately 36 percent of sampled students attend public schools, which receive public subsidies, are managed by local municipalities or Local Educational Services, and enroll students mainly from low-income families. The remaining students attend private (non-subsidized) schools, which are funded entirely by tuition fees and serve affluent families. Table 2 indicates that the sample is balanced across nearly every observed dimension.

To verify the success of our randomization strategy, we conducted balance tests following Kerwin et al. (2024). Specifically, we performed an omnibus F-test of joint orthogonality with randomization inference. We conducted 500 simulations, randomly permuting treatment assignment within strata and computing F-statistics for joint significance of baseline covariates. The observed F-statistic from our actual treatment assignment (0.155) falls below the 5th percentile of the simulated distribution (0.156), indicating that covariates are well-balanced across treatment arms. Additionally, we ensured balance not only across treatment arms but also across implementation weeks, as students were contacted in different weekly waves.

3 Hypotheses and Data

Building on ([Ajzenman et al., 2024](#)), who documented positive effects of human mentors but limited impact of rule-based chatbots on attracting students to teacher education programs in Chile, we formulate the following hypotheses about the relative effectiveness of human mentors versus AI-powered chatbots:

A1 (*Human mentor effectiveness*). Students assigned to the human mentor treatment will be more likely to rank a teacher education program as their first choice compared to the control group. In the 2021 experiment, human mentors increased this likelihood by 1.25 percentage points from a baseline of 13.1%.

A2 (*AI chatbot effectiveness*). Students assigned to the AI chatbot treatment will show a higher probability of ranking a teacher education program as their first choice compared to the control group. We expect larger effects than the rule-based chatbots used in 2021 (which showed positive but insignificant effects), given the enhanced interaction capabilities of generative AI.

A3 (*Treatment differential*). Although we do not have a strong prior, our expectation is that the technological advances of the smart chatbots will improve their performance versus the [Ajzenman et al. \(2024\)](#)’s chatbots. This expectation is supported by recent evidence showing that large language models can achieve performance comparable to human experts in complex tasks ([Goh et al., 2024](#)). However, it is not clear that the improvement will be strong enough to lead to a significant and economically meaningful effect.

Our analysis relies on two main data sources. The first is administrative data from DEMRE, encompassing rich individual-level information about students. This includes demographic characteristics (gender, age), educational background (type of high school attended), and academic performance indicators (high school GPA and PAES scores). Crucially, DEMRE also provides detailed information about the university application process, including students’ ranked preferences for university-major combinations, their admissions results, and final enrollment decisions. The data also contains information about students’ socioeconomic status and parents’ educational attainment, though

this information is self-reported during the PAES registration process.

The second source consists of detailed interaction data from WhatsApp conversations between students and both human mentors and the chatbot. This data captures the nature and intensity of the engagement, including response rates, conversation duration, and types of questions asked. For the human mentoring arm, we observe how mentors adapt their responses to students' specific concerns and interests. For the chatbot arm, we can analyze how the AI system interprets and responds to student queries, providing insights into the effectiveness of different communication strategies. All interaction data is automatically generated and anonymized, with students identified only through the masked IDs provided by DEMRE.

Our analysis examines four outcomes regarding student choices. The primary outcome is whether a student ranked an education major as their first choice, which indicates a strong preference for pursuing an education career. We also analyze the proportion of education majors in a student's choice set, and whether they applied to at least one education major as primary outcomes.

We also consider actual enrollment in education programs as a secondary outcome, indicated by enrollment in an education major program. This outcome should be interpreted cautiously. This is because enrollment is determined by a centralized assignment algorithm which considers the choices of students in any arm (treatment or control); thus, the choice of a treated student affected by the intervention could affect the enrollment outcome of a control student, biasing the estimation of that specific outcome. This potential problem could affect the enrollment outcome, but not the main outcomes related to students' major choices.

4 Research Plan

4.1 Identification and Estimation

We employ three main specifications in our empirical analysis. The first two specifications use the following OLS regression:

$$y_i = \alpha T_i + X_i\beta + \epsilon$$

where y_i represents the outcome variables for student i (described in section 3), and T_i represents the treatment assignment. In our primary specification, T_i is a vector of two dummy variables indicating assignment to either human mentoring or chatbot treatment, allowing for pairwise comparisons between all experimental arms (human mentoring versus control, chatbot versus control, and human mentoring versus chatbot). In our secondary specification, T_i is a single dummy variable that equals one if the student was assigned to either treatment arm, estimating the pooled effect of any form of mentoring. X_i is a vector including the stratification variables: gender, type of high school (public, private-subsidized, or private non-subsidized), and initial career interests (education or social sciences).

For our third specification, which focuses only on the treated groups (humans versus chatbots), we will estimate the Local Average Treatment Effect (LATE) of message engagement:

$$y_i = \alpha E_i + X_i\beta + \epsilon$$

where E_i represents the student's level of engagement measured by the number of messages exchanged between the student and their assigned mentor (human or chatbot). We will instrument E_i with the random assignment to human versus chatbot mentoring. This specification allows us to estimate the causal effect of message intensity while accounting for the endogenous nature of student responses.

For all specifications, we will use robust standard errors as randomization was conducted at the individual level with no clustering. Our main analysis follows an intention-to-treat approach, though the LATE specification provides insight into the role

of actual engagement with the interventions.

4.2 Inference and Power

We base our sample size calculation on the results from [Ajzenman et al. \(2024\)](#). The previous study used the same context, population, and similar interventions. [Ajzenman et al. \(2024\)](#) found a 1.25 percentage point increase in the probability that a student ranks an education program as their first choice in the humans arm (mean in the control group was 13%). Considering that outcome, assuming no clustering, power of 0.8, a sample size of 9,000 individuals in the humans arm and 14,500 in the chatbot arm would allow us to identify an effect of approximately 1.25 pp (at 5%) using demographic controls. Similarly, a sample size of 20,000 individuals in the chatbot arm and 14,500 in the control arm would allow us to identify an effect of 1 pp. Finally, we would be able to identify differences of approximately 1.3 pp between the two treatment arms.

4.3 Heterogeneity

Our primary hypotheses build on findings from the previous version of this intervention ([Ajzenman et al., 2024](#)), which revealed significant heterogeneity in treatment effects. As in ([Ajzenman et al., 2024](#)), we plan to study heterogeneity using causal forest techniques ([Athey et al., 2019](#)). We expect to find similar patterns as in ([Ajzenman et al., 2024](#)), where the analysis uncovered substantial heterogeneity in treatment effects, mainly in terms of gender, student performance, and to a lesser extent socioeconomic status. Specifically, higher-performing (in terms of high school GPA and scores on the college entrance exam) and male students consistently experienced the most significant treatment effects across most outcomes.

Additionally, we will analyze heterogeneous effects based on students' declared area of interest (education versus social sciences), as the effectiveness of both human tutors and AI chatbots may vary depending on students' initial career preferences. This analysis is particularly relevant as it will help us understand whether the interventions are more effective at reinforcing existing interest in teaching careers or at attracting students from related fields to consider education programs.

References

- Ajzenman, N., Elacqua, G., Hincapié, D., Jaimovich, A., Boo, F. L., Paredes, D. and Román, A. (2021). Career choice motivation using behavioral strategies, *Economics of Education Review* **84**.
- Ajzenman, N., Elacqua, G., Jaimovich, A. and Pérez-Núñez, G. (2024). Humans versus chatbots: Scaling-up behavioral interventions to reduce teacher shortages, *Journal of Political Economy Microeconomics* . Forthcoming.
- Athey, S., Tibshirani, J. and Wager, S. (2019). Generalized random forests, *The Annals of Statistics* **47**(2): 1148–1178.
- Elige Educar (2020). Análisis y proyección de la dotación docente en contextos rurales.
- Goh, E., Gallo, R., Hom, J. et al. (2024). Large language model influence on diagnostic reasoning: A randomized clinical trial, *JAMA Network Open* **7**(10): e2440969.
- Kerwin, J. T., Rostom, N. and Sterck, O. (2024). Striking the right balance: Why standard balance tests over-reject the null, and how to fix it, *IZA Discussion Paper* (17217).
- Medeiros, M. P., Gómez, C., Sánchez, M. J. and Orrego, V. (2018). Idoneidad disciplinar de los profesores y mercado de horas docentes en Chile, *Calidad en la Educación* **48**: 50–95.

Appendix A: Initial Messages

A.1 Human mentor

[Elige Educar mentor]: Hi, [student name]! My name is [mentor name], I'm a pedagogy student in [discipline] and a mentor of the "Quiero Ser Profe" program from Elige Educar.

When you registered for the PAES exam, you agreed to receive information to guide your career decision. That's why I'm offering you a WhatsApp mentoring session to answer your questions about education degrees.

This conversation will only be open for 24 hours, so take advantage of talking to me to get all your questions answered!

Choosing a career is not easy and I'm here to guide you. Would you like to know how I can help you? Choose an option!

[Button 1: Questions?]

[Button 2: Let's get started ;)]

[Button 3: I don't want to participate]

[Button 1: Questions?]

[Elige Educar mentor]: I can tell you about my experience, answer your questions about pedagogy careers and guide you through the university application process! Some topics I can help you with:

- Requirements to apply for pedagogy
- NEM or Ranking
- Scholarships
- Universities and accreditation
- Future Employment

- Vocational guidance: How to know if pedagogy is right for you

To offer you personalized information, I would like to know more about you. Can you tell me what region you are from, what your career of interest is and if you are in 12th grade or have already graduated?

[Button 2: Let's get started!]

[Elige Educar mentor]: Great!

Have you already decided to study pedagogy, or do you still not know if it's for you? Choosing a career is not easy and I'm here to guide you and tell you about my experience. I can inform you about requirements, scholarships, or even how to choose a university. Tell me what doubts you have and what you would like to know about the pedagogy career.

[Button 3: I don't want to participate]

[Elige Educar mentor]: I understand, [student name]! If you ever need information, you can visit the "Quiero Ser Profe" program site on our website here: <https://eligeeducar.cl/quieres-ser-profe/>. Good luck!

A.2 Chatbot Kai

[Kai]: Hi [student name]! This is Kai, your Elige Educar mentor powered by artificial intelligence.

When you registered for the PAES exam, you agreed to receive guidance from Elige Educar. Our "Quiero Ser Profe" program offers you free WhatsApp mentoring to support you in your career decision. I was created together with a team of professionals for this purpose.

Choosing a career is not easy and I'm here to guide you. Would you like to know how I can help you? Pick an option!

[Button 1: Let's get started!]

[Button 2: I don't want to participate]

[Button 3: How do I work?]

[Button 4: What can I help you with?]

[Button 1: Let's get started!]

[Kai]: Great! Have you already decided to study education, or are you still unsure if it's right for you? Choosing a career is not easy and I'm here to guide you. I can inform you about requirements, scholarships, or even how to choose a university. Tell me what questions you have and what you would like to know about education degrees.

[Button 2: I don't want to participate]

[Kai]: I understand, [student name]! Don't worry, if at any point you need information, you can visit the "Quiero Ser Profe" program site on our webpage here: <https://eligeeducar.cl/quieres-ser-profe/>. I wish you much success on your journey!

[Button 3: How do I work?]

[Kai]: I'm here to answer all your questions about education degrees, guide you in your career decision, and offer you the necessary resources so that you can make the best

choice. I was developed by the Elige Educar team, especially to support you. I'll be available today and you can write to me whenever you want until January 9th. Also, I will contact you during the application process (January 6-9) to provide support.

I can answer questions related to requirements, scholarships, or even how to choose a university. I'm here to guide you.

Tell me: Have you thought about being a teacher? What doubts do you have about studying education?

[Button 4: What can I help you with?] [Kai]: I'm here to help you answer your questions about education degrees and guide you through the university application process! Here are some topics that might be of interest to you:

- What are the requirements to apply for an education degree?
- What is NEM? What is ranking?
- What are the scholarships for higher education, and how do you apply?
- What should I look for when choosing a university?
- Are there jobs in education? What is the work like?
- How do I know if education is right for me?

What questions do you have about education degrees and the application process? Let me know!