WORKING PAPER Nº IDB-WP-01501

Humans versus Chatbots: Scaling-up behavioral interventions to reduce teacher shortages

Nicolás Ajzenman
Gregory Elacqua
Analia Jaimovich
Graciela Pérez-Nuñez

Inter-American Development Bank
Education Division

July 2023

Humans versus Chatbots: Scaling-up behavioral interventions to reduce teacher shortages

Nicolás Ajzenman
Gregory Elacqua
Analia Jaimovich
Graciela Pérez-Nuñez

Inter-American Development Bank
Education Division

July 2023

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Humans versus bots: scaling-up behavioral interventions to attract more and better applicants into teacher education programs / Nicolás Ajzenman, Gregory Elacqua, Analía Jaimovich, Graciela Pérez-Núñez.
p. cm. — (IDB Working Paper Series ; 1501)
Includes bibliographical references.
1. Education-Study and teaching-Chile. 2. Teachers-Selection and appointment-Chile. 3. Teachers-Supply and demand-Chile. 4. Education-Effect of technological innovations on-Chile. I. Ajzenman, Nicolás. II. Elacqua, Gregory M., 1972- III. Jaimovich. Analía. IV. Pérez-Nuñez, Graciela. V. Inter-American Development Bank. Education Division. VI. Series.
IDB-WP-1501

JEL Codes: D91, I23, I25
Keywords: Teachers, teacher policy, teacher shortages, scale-up, behavioral, bots

http://www.iadb.org

Humans versus Chatbots: Scaling-up behavioral interventions to reduce teacher shortages*

Nicolás Ajzenman† Gregory Elacqua‡ Analía Jaimovich §
Graciela Pérez-Núñez ¶

July 2023

Abstract

Empirical results in economics often stem from success in controlled experimental settings, but often fail when scaled up. This study presents a behavioral intervention and a scalable equivalent aimed at reducing teacher shortages by motivating high school students to pursue an education degree. The intervention was delivered through WhatsApp chats by trained human promoters (*humans* arm) and rule-based Chatbots programmed to closely replicate the *humans* program (*bots* arm). Results show that the *humans* arm successfully increased high-school students' demand for and enrollment in education majors, particularly among high-performing students. The *bots* arm showed positive but smaller and statistically insignificant effects. These findings indicate that a relatively low-cost intervention can effectively reduce teacher shortages, but scaling up such interventions may have limitations. Therefore, testing scalable solutions during the design stage of experiments is crucial.

***JEL classification*: D91,I23, I25**

Keywords: *Teachers, teacher policy, teacher shortages, scale-up, behavioral, bots*

*Trial registration number AEARCTR-0008269. We thank Elige Educar, the Departamento de Evaluación, Medición y Registro Educacional (DEMRE) of the University of Chile, and the Centro de Estudios of the Chilean Ministry of Education for providing the data employed in this study. We thank Ruben Durante, Bruno Ferman and Fernanda Estevan for valuable comments. We also thank Karina Córdoba, Gonzalo Escalona, and Íbes Berríos for their valuable contribution in designing and implementing the interventions. We also thank Diana Hincapié for her contribution to the design and funding of the experiment and research. Finally, we thank Matias Méndez for his diligent work supporting the data analyses. We gratefully acknowledge the Inter-American Development Bank for funding the research presented herein. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent. The authors have no conflicts of interest or financial or material interests in the results. All errors are our own.

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

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

§Inter-American Development Bank. E-mail: analiaj@iadb.org.

¶Inter-American Development Bank. E-mail: gracielap@iadb.org.

1 Introduction

Since the credibility revolution and the incorporation of experiments as one of the mainstream empirical tools in economics, reliable, internally valid policy evaluations have grown massively. These empirical results helped nurture the formulation and implementation of interventions in different fields and sectors, which increased their policy relevance and impact (Angrist and Pischke, 2010). The widespread adoption of experiments has created a new concern: scalability (Banerjee et al., 2017). As noted by Al-Ubaydli et al. (2020) and List (2022), many results that worked in experimental settings yield much smaller effects when policymakers try to scale them up. This problem is especially troublesome considering the extent to which public policy relies on results from experimental settings.

One of the roots of scaling-up failures is the "representativeness of the situation" in experimental settings (List, 2022; Al-Ubaydli, List and Suskind, 2017). Experimental studies are typically administered in exceptional situations with intensive oversight. When programs are scaled up, implementation details escape researcher control, and protocol adherence decreases. Thus, implementation and delivery problems are more likely to arise. Technology is a promising avenue (Al-Ubaydli et al., 2021) to promote standardization, ensure correct dosages, and, more broadly, minimize effect size losses at scale. However, in most cases, the evaluation of specific pilots is not matched with a suitable (and evaluated) plan to scale it up, thus making it hard to anticipate expected effect size losses. This paper presents the results of a pilot program and a scalable equivalent in the same experiment. The scalable equivalent was implemented using a widely used technology: rule-based Chatbots,[1] designed to standardize treatment and reduce the probability of incorrect delivery. We purposely chose the simplest type of chatbot to test a technology that is very inexpensive and easy to implement with minimal technical knowledge.

Our experiment focuses on a prototypical intervention based on insights from behavioral economics: a low-cost policy to reduce teacher shortages by motivating high school

[1] Rule-based Chatbots, unlike their AI-based equivalents, are not trained and do not learn. They are pre-programmed using fairly complex decision trees. These are arguably the most common Chatbots as they are inexpensive and easy to implement. See, for instance, Abd-Alrazaq et al. (2020).

students to pursue an education degree (Ajzenman et al., 2021). The standard intervention (*humans*) was delivered by trained human promoters through WhatsApp chats and, although it is a relatively low-cost intervention, could have potentially hidden scaling-up costs. The scalable equivalent (*bots*) was delivered by WhatsApp rule-based Chatbots carefully programmed to replicate every dimension of the *humans* original program as much as possible. The scripts in both cases were based on objective information about the higher education application process and behavioral insights emphasizing intrinsic, extrinsic, prosocial, and prestige-based motivations.

The interventions aimed to address the shortage of qualified teachers in Chile, projected to reach over 30,000 by 2025, accounting for 12 percent of the total teacher supply (Medeiros et al., 2018). This shortage is a pressing problem that affects many developing countries (Elacqua et al., 2022b), and the interventions were specifically designed to attract high-performing high school students into the teaching profession. The experiment was conducted with 40,813 final-year high-school students who declared on a survey conducted by the central higher education testing authority (DEMRE, *Departamento de Evaluación, Medición y Registro Educacional*) that they might be interested in pursuing a career in education or, more generally, the social sciences. Students, who explicitly consented to be contacted to receive college major-related information, were randomly assigned to one of the two treatment arms (*humans* or *bots*) or the control group.

The messages were delivered to students before they applied to college majors. Chile uses a nationwide centralized admission system that assigns students to university-major combinations based on their preferences and academic performance. Students can apply to up to ten combinations of university majors, and admission is granted based on a weighted average of their high school GPA and scores on the university entrance exam.

The *humans* program was successful: high-school students were 1.3 percentage points more likely to list an education major as their first choice (control mean: 13%) and 0.9 percentage points more likely to enroll in an education major (control mean: 10%).[2] This effect, if scaled, would represent a reduction of between 0.8 to 2 percent of the teacher

[2] As we explain further in the paper, the enrollment outcomes should be interpreted cautiously. This is because the effect on the choice of the treatment group could have affected the enrollment of students in the control group.

deficit projected in Chile by 2025. In contrast, although the *bots* program showed positive effects, the effect sizes were between two-thirds and one-half smaller and generally insignificant.

Students in both arms received several messages, mostly through WhatsApp, to motivate them to major in education. The scripts of the two treatment arms (*humans* and *bots*) were carefully designed to be as similar as possible by members of *Elige Educar*, a Chilean NGO co-responsible for implementing the experiment whose mission is to improve the status of the teaching profession and attract higher-performing high school students into education. The messages were intended to motivate students to pursue a degree in education by appealing to four types of potential motivations often emphasized by the literature (Hendricks, 2014; Hoyle, 2001; Watt et al., 2012): **intrinsic**, related to enjoying the tasks of the job; **altruistic**, related to how teachers can make a difference in the world; **extrinsic**, related to the material benefits of having a full-time teaching position, such as a low unemployment rate, longer paid vacations, an above-median salary; and **prestige**, related to the perceived social value of teachers. Individuals in the control group did not receive any messages.

The messages in the *humans* arm were delivered by forty-three tutors recruited from teacher education programs and trained by *Elige Educar*. Tutors followed a pre-designed script, which included potential answers to student questions. The tutors contacted 8,161 individuals who took the college entrance exam (primarily high school students). The team of tutors began sending messages the last week of November 2021 and ended the process by the second week of January 2022. In the initial contact, tutors offered the option of a phone call or a conversation via WhatsApp, and only a small number of students chose a phone call. Tutors had a contact protocol consisting of three attempts every five days to complete a conversation. If a student did not answer, they received a summary of the remaining information in the script. Students in any arm could request to stop receiving messages or calls at any point. The *humans* arm required significant human resources as tutors needed to be trained and their work involved in-depth interactions with students over several sessions.

Messages in the *bots* arm were delivered by rule-based Chatbots, which required much less human resources than the *humans* arm. The script was also pre-designed, closely following the one used in the *humans* arm (see Appendix C). The *bots* and the human tutors were able to answer specific questions with standardized answers, based on the experience of *Elige Educar*. *Bots* sent messages from the first week of December 2021 to the second week of January 2022.

Both the *humans* and *bots* delivered information on topics such as motivation (e.g., how teachers can positively impact the lives of thousands of students), the economic reality of being a teacher (including expected salaries and paid vacation days), and the prestige of teaching careers (based on factual information from surveys about perceptions of different career options). The information provided was tailored to student questions and answers in both treatment arms. For instance, if the primary concern of a student was related to financial matters, the *humans* or *bots* would provide specific information on that topic. The goal was to provide information that would be helpful to students based on their motivations and to avoid providing information that could potentially backfire, as seen in other studies in a similar context (Ajzenman et al., 2021).

To further explore the potential policy implications of our interventions, we use causal forest techniques to examine any heterogeneous treatment effects and identify student characteristics that maximize the effect of the *humans* arm using an honest approach (Athey and Imbens, 2017). Our analysis uncovers substantial heterogeneity in treatment effects, mainly in terms of gender, student performance, and to a lesser extent socioeconomic status. In particular, higher-performing (in terms of GPA and scores on the college entrance exam) and male students consistently experienced the most significant treatment effects across most outcomes. Students from lower-income families and non-private schools (i.e., public and subsidized-private schools) also responded more. The fact that the intervention mainly benefited high-performing students could be significant for policy since high-quality teachers are critical for student learning, particularly for students from lower-income or disadvantaged backgrounds (Lankford et al., 2002).

Our paper contributes to two strands of the literature. First, it relates to the emerging literature on scaling up experimental evidence (Muralidharan and Niehaus, 2017; Vivalt, 2020; DellaVigna and Linos, 2022). Several papers, such as Al-Ubaydli, List, LoRe and Suskind (2017); Al-Ubaydli et al. (2020, 2021), have elaborated on several threats that experimental projects can face when scaled up and, in some cases, propose more effective designs to overcome scale issues. We contribute to this literature by testing the effects of a program and a scalable counterpart in the same experiment, using technology (Chatbots) to standardize the treatment, thus reducing the probability of incorrect delivery. As far as we are aware, this paper presents one of the first experiments in which a program is simultaneously designed and tested with a scalable equivalent.[3]

Second, our paper contributes to the literature on teacher supply (Hoxby and Leigh, 2004; Corcoran et al., 2004; Boyd et al., 2006; Bacolod, 2007; Boyd et al., 2013) and career motivation (Han et al., 2018). While much of the literature has focused on the effects of monetary incentives, such as salary increases, on teacher supply and retention (see, for instance, Clotfelter et al. (2008b); Falch (2011); Baron (2021); De Ree et al. (2018)), we explore the effectiveness of a non-monetary intervention to encourage high-performing high school students to pursue a career in education. Although emerging evidence suggests that students' motivations to pursue careers in education are influenced by various factors such as extrinsic, intrinsic, and prestige concerns (see Han et al. (2018); Perryman and Calvert (2020)), there is limited research on the teacher supply impact of interventions based on these motivations (Ajzenman et al., 2021). Our paper contributes to this literature by providing evidence on how a low-cost intervention that utilizes behaviorally-informed messages appealing to different motivations affects teacher supply.

The potential policy implications of these results are significant. Firstly, the evaluated strategies address a crucial issue in education and development: teacher supply shortages. Teachers are critical inputs in the education production function and significantly impact student performance (Rivkin et al., 2005; Jackson, 2018). Therefore, shortages can have detrimental effects, particularly in vulnerable areas (Ajzenman et al., 2023, 2022), contributing to widening socioeconomic achievement gaps. Secondly, by explicitly comparing the impact of the standard program and a scalable equivalent, our results shed light on the limitations of such interventions.

[3] A small but growing literature related to this paper uses Chatbot technology to implement interventions in education. Examples of this in different contexts include Page and Gehlbach (2017) and Nurshatayeva et al. (2021). Our intervention uses a simpler Chatbot.

2 Chilean context

Teachers are a crucial input in the educational production function, having the most significant impact on student learning and accounting for the highest share of educational budgets (Rockoff, 2004; Rivkin et al., 2005; Staiger and Rockoff, 2010; Chetty et al., 2011, 2014). However, motivating high-performing students to become teachers is a challenge (Bertoni et al., 2020; Elacqua et al., 2022b). While many countries have introduced policies to attract good teachers to regions where they are most needed (Clotfelter et al., 2008a; Jackson, 2009; Ajzenman et al., 2022, 2023; Elacqua et al., 2022a; Pérez-Núnez, 2020), interventions to motivate students to become teachers are scarcer and often not very effective (Ajzenman et al., 2021). Like many countries in Latin America, Chile faces a severe shortage of quality teachers in its school system. Estimates suggest that by 2025 there could be a scarcity of more than 30,000 qualified teachers (Medeiros et al., 2018).

Over the last decade, various Chilean governments have adopted policies to improve teacher quality and attract better students to teacher education programs. In 2012, the Ministry of Education relaunched a full-tuition scholarship for high-performing applicants to teaching programs (Pérez-Núnez, 2020). Further, between 2014 and 2018 the Chilean government implemented major reforms across all education levels, from preschool to higher education, as well as significant improvements to teachers' professional careers. One key reform was the provision of free tuition for postsecondary studies in education to students in the bottom 50 percent of the income distribution, which was designed to automatically extend the benefit to additional income segments as tax revenues reached specific thresholds. Since 2018, the policy has been expanded to cover students in the bottom 60 percent of the income distribution. Additionally, some teaching regulations (Ley 20.903, *Sistema de Desarrollo Profesional Docente*) have been updated to improve working conditions for teachers and attract better candidates to the teaching profession. This reform was designed to be gradually implemented from 2017 to 2025 and established a new teacher career ladder with stages and progression criteria, including higher admission requirements for teacher education programs (measured by scores on the college entrance exam), quality assurance for teacher education, support for new teachers, and improved working conditions such an increase in teacher salaries of more than 30 percent, a reduction in the number of hours of instruction, and more hours for course preparation, performance evaluations, and school-focused professional development (Ávalos and Bellei, 2019).

While all these reforms aimed to enhance the overall quality of education, they had unintended short-term effects on enrollment in teacher education programs. For example, the free tuition policy decreased demand for teaching programs among high-achieving students, particularly those from low-income backgrounds who could opt for majors with higher economic returns (Castro-Zarzur et al., 2022). Additionally, the new legal requirement of higher entry scores for teacher education programs reduced the pool of eligible candidates, decreasing total enrollment.

Chile's higher education system comprises universities, professional institutes, and technical vocational schools. Only certified universities offer teacher education programs. Students apply to universities through a nationwide centralized admission system that allocates them to university-major combinations based on their preference rankings and academic performance on the university entrance exam. The exam was administered in the second week of December 2021 and results were announced by mid-January 2022. The entrance exam is a standardized test covering several subjects organized and conducted by the Department of Evaluation, Measurement and Educational Registration (DEMRE) that students must take to apply to the vast majority of Chilean higher education programs. Students rank up to ten combinations of universities and majors. Admission is granted based on a weighted average of high school GPA and scores on different sections of the university entrance exam, including the mandatory verbal and math tests and specific tests depending on major.

Of the 275,631 students who registered for the exam, 85 percent took the mandatory verbal and math tests, which made them eligible to apply to university majors. After the results were announced, the higher education application period began, allowing students to rank up to ten major-college preferences. Programs ranked applicants based on a weighted score that included math, verbal, GPA, and social sciences or science scores. Each university-major has different admissions requirements. For example, teacher education programs require a minimum score of 500 points (on a 150-850 scale in 2021). Using student rankings and program-specific scores, the system assigned each student to one program utilizing the Gale and Shapley student-proposing deferred acceptance algorithm (Gale and Shapley, 1962), ensuring only stable matches. Admission results were published by the end of January, and students completed the necessary paperwork at their admitting institution to finalize their enrollment.

2.1 The program

Elige Educar, a Chilean NGO founded in 2009, was in charge of implementing the large-scale intervention that we evaluate in this paper in 2021. Elige Educar's mission is to ensure all children have quality teachers. They focus on three goals: (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. To attract students to education majors, since 2010 Elige Educar has focused on two main initiatives: the "*Quiero ser Profe*" program and a massive media campaign. Most Elige Educar campaigns are conducted in partnership with the Chilean Ministry of Education.

The *Quiero Ser Profe* ("I want to be a teacher") program offers individual tutoring by phone to students interested in teacher education programs a few months before they apply to higher education institutions. Tutors are typically university students majoring in education who plan to become teachers. In addition to individual tutoring, 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.

When registering for the university entrance exam, students must complete a survey administered by the Department of Evaluation, Measurement, and Educational Registration (DEMRE), which oversees the centralized application process. The survey asks about students' interests (for instance, which fields and majors they are considering). In 2021, the survey asked students whether they consent to DEMRE sharing their contact information with Elige Educar to provide counseling assistance with their higher education application process (see Appendix B). Of the 275,631 students who registered for the entrance exam, 64.3 percent agreed to share their contact information with Elige Educar (177,224). Our sample for the experiment consisted of 40,813 students who expressed an interest in education majors or, more generally, social science majors, as these preferences were deemed more receptive to our interventions aimed at promoting teacher education programs. The interventions were implemented during the university application process, starting before the university entrance exam date and continuing until just before the application deadline.

3 Experimental design

The randomized controlled trial (RCT) was carried out with 40,813 students registered to take the university entrance exam who declared an interest in pursuing a career in teaching or social sciences on a DEMRE survey (see survey questions in Appendix B). Of this sample, the 24,487 applicants who expressed an interest in education were randomly assigned to the *humans* treatment (8,161), the *bots* treatment (8,163), or the control group (8,163). A sub-sample of 16,326 individuals who expressed an interest in social sciences (but not teaching) were randomly assigned to the *bots* treatment (8,163) or the control group (8,163).

The sample was stratified by gender (male, female), type of high school they attended (public, private-subsidized, or private), and high school academic performance (top 30%, bottom 70%). For more details on the stratification, refer to Table A1 in the Appendix.

The decision to separate between those who expressed explicit interest in an education career versus a social sciences career was driven by resource constraints. Elige Educar had the capacity to train mentors to reach approximately 8,000 students, which set a limit on the number of individuals that could be assigned to the *humans* arm. Conversely, there were no constraints on the number of individuals that could be reached through the Chatbot. As Elige Educar's mentoring program is designed for students interested in education, we allocated all the mentors to students with a specific interest in education majors. To avoid losing the sample interested in social sciences, all such students were allocated to the *bots* arm or the control group. Therefore, the treatment effects from the *humans* arm are valid only for students interested in an education major, whereas the treatment effects from the *bots* arm are valid for those interested in both education and social sciences majors.

In the main specifications, we report the results using the entire sample and include a dummy controlling for interest (teaching or social sciences). In the Appendix, we report the effects considering only students interested in education. Given that the results of the *humans* arm on students interested in education and social sciences are not significantly different, the results in the main tables are comparable across arms.

3.1 Experimental arms

As previously mentioned, the *humans* treatment arm is a text message-based intervention carried out by trained tutors who initiate conversations via WhatsApp and offer a phone call option. While tutors were instructed to follow a script, they had some flexibility to respond to individual questions and interactions. In contrast, the *bots* treatment arm relies on a Chatbot that initiates a pre-programmed conversation via WhatsApp using a decision tree. The *bots* followed a script that was purposely similar to the one used by the human tutors, although with less flexibility. Chatbots were rule-based, meaning they were not AI-based like those used in other education studies (e.g., Page and Gehlbach (2017); Nurshatayeva et al. (2021)). Although rule-based Chatbots are less intelligent, they are more affordable and easier to implement. We chose the simplest technology possible to make programming the Chatbots accessible to those with minimal technical skills.

The human tutors contacted the sampled individuals from the last week of October 2021 until the second week of January 2022 (11 weeks), while Chatbots sent messages to different groups of sample individuals from the beginning of November 2021 until mid-January 2022 (7 weeks). The control group did not receive any messages. The difference in treatment length does not affect treatment intensity but rather reflects the productivity of each method. Indeed, the *bots* arm could have been completed immediately but was spread out throughout different weeks.

The scripts for both treatment arms aimed to inform high school students about the college application process and encourage them to pursue an education degree. The scripts were designed to make salient the four types of motivations typically emphasized by the literature on job preferences as the primary motivating drivers for individuals pursuing professional careers and, specifically, education careers (Skatova and Ferguson, 2014; Watt et al., 2012; Hendricks, 2014; Hoyle, 2001). First, intrinsic motivation, understood as enjoying the job's tasks. Second, extrinsic motivation, understood as valuing the material working conditions. Third, prestige-based motivation, understood as placing importance on the career's societal status. Finally, altruistic motivation, which refers to enjoying helping others (this could be considered a specific type of intrinsic motivation). We show the structure of the scripts in Appendix C.

There are some key aspects of each script's design. The scripts were designed to be conversational rather than a series of identical messages, allowing for customization based on individual student responses (see Appendix C). This approach was informed by insights from Ajzenman et al. (2021), who found a null (or even backfiring effect) of a three-arm email campaign which made three types of motivations salient: intrinsic/altruistic, extrinsic, and prestige. The results showed that emphasizing intrinsic and prestige factors reduced the number of high-achieving students applying to education majors while highlighting extrinsic rewards increased applications among lower-performing students. In Chile, high-performing students typically come from privileged backgrounds. Therefore, emphasizing the intrinsic and prestige values of teaching could draw attention to the social status disparity between education careers and other professions (such as doctors or lawyers). This may discourage these students from pursuing education majors, as suggested by the authors. Conversely, emphasizing the improved economic prospects of teaching appealed to lower-performing students, who are more likely to come from disadvantaged families and may place a higher value on financial rewards. The key lesson from the study is that emphasizing different motivations may lead to unexpected outcomes if not tailored to the appropriate audience. To avoid unintended outcomes, we designed the scripts to allow students to reveal their motivations beforehand and tailored the messages accordingly. While motivations are not completely independent, tutors and Chatbots were instructed to emphasize each student's most suitable motivation(s).

For instance, if a student expressed interest in pursuing an education major, the tutors would ask about the primary characteristics of teaching that motivate them. If the student mentioned social impact as a motivator, the tutors would emphasize that "an education career can make you an agent of change. Did you know that throughout their professional career, a teacher can have a positive impact on the lives and opportunities of up to 5,000 children?". If a student expressed interest in a specific topic and a desire to pursue a degree that would allow them to learn and work with those concepts, the tutor would respond by saying, "Teaching is an excellent choice if you enjoy X (a specific discipline) and want to share your passion with others. This is especially true if you have skills in teaching and working with children and/or young people." If a student expressed interest in material or monetary incentives, the tutor would answer with "Are you familiar with the new Teachers' Policy? It is a law that has improved teachers' working conditions since 2017. The policy has increased salaries by 30%, with ongoing increases based on experience and performance. Also, 35% of total work time is now reserved for class preparation, providing teachers with a better work-life balance." If a student expressed reluctance to pursue an education career due to concerns about prestige, the tutor would respond "Are you aware that education is one of the four most highly valued professions in Chile, alongside medicine, engineering, and law? Also, did you know that education, medicine, and dentistry are the only degrees that can be taught exclusively in accredited institutions? Pursuing education can lead to a highly respected and prestigious career in society!"

In addition to addressing motivations, the messages also included general information about the application process and education majors. Although the scripts for the two arms -tutors and Chatbot- were not identical due to practical limitations, we tried to write each script comparably while considering the structured nature of Chatbot conversations. Each script began with a brief introduction, followed by a section on career motivations and concerns, a segment on general career information, and a brief module on the application process. Further information can be found in Appendix A.

3.2 Data: Descriptive statistics and balance

We identified 177,224 students who met the eligibility criteria for the study: those who were enrolled to take the university entrance exam, agreed to share their information and be contacted by Elige Educar (for counseling about the higher education application process and research purposes), and expressed interest in pursuing a major in education or social sciences. From this pool, we selected a sample of all 40,813 students who indicated on the survey that they were potentially interested in pursuing a major in education or social sciences (see Appendix B). The remaining students (approximately 136,411) were excluded because they were part of a separate study. Our analysis was conducted on the sub-sample of individuals who provided valid contact information and who received at least the first WhatsApp message, resulting in a sample size of 39,119, or 96 percent of the original sample.

We also use anonymous administrative data from DEMRE, the Chilean agency responsible for university admissions. The data comprises individual-level information on various demographic and educational characteristics of students, such as gender, year of birth, year of high school graduation, type of high school attended (public, private-subsidized, and private), high school GPA, university entrance exam scores, and two variables related to their parents: socioeconomic status (defined as family income below or above the poverty line) and educational attainment. Table 1 presents descriptive statistics of the covariates used in our analysis. The data also includes information on applications, admissions, and enrollment in universities that use the centralized application system. To supplement this data, we obtained official records from the Ministry of Education on enrollment for all higher education institutions in Chile.

Table 1: Baseline covariates summary statistics

	Obs.	Mean	Std. Dev.	Min	Max
Female	39,005	0.665	0.472	0	1
High school type					
Public	38,501	0.345	0.475	0	1
Private-Subsidized	38,501	0.577	0.494	0	1
Private	38,501	0.078	0.268	0	1
High school GPA, top 30%	39,005	0.337	0.473	0	1
Recent high school graduation	39,005	0.809	0.393	0	1
High school GPA	38,577	5.835	0.496	4	7
High school ranking score	38,577	608.4	125.1	211	850
Average math and verbal test	33,562	493.1	91.1	150	822
Parents without high school education	37,578	0.203	0.402	0	1
Parents with high school education	37,578	0.588	0.492	0	1
Parents with higher education	37,578	0.209	0.406	0	1
Low-income family	33,902	0.628	0.483	0	1

The sample is composed of 67 percent women, with 58 percent of students attending private-subsidized schools, primarily owned and operated privately but receiving per-student public subsidies and catering mainly to middle-class families. Approximately 35 percent of sampled students attend public schools, which receive public subsidies, are managed by local municipalities, and enroll students mainly from low-income families. The remaining students attend private (non-subsidized) schools funded entirely by tuition fees and serve affluent families.

Additionally, 81 percent of sampled students graduated from high school in the year before or the year of the 2021 university entrance exam, and 63 percent of students are from families with (self-reported) incomes below the poverty line. The self-reporting here is an important caveat; many students do not answer this question, but it is still useful to test for balance across treatment arms. Tables 2 and 3 indicate that the final sample is balanced across nearly every observed dimension. Although there may be a plausible concern that the intervention impacted not only choices but also performance on the entrance exam, the table shows that exam scores are almost identical across all groups, which reduces this concern.[4]

[4]Furthermore, the proportion of students who took the exam is practically the same across treatment and control arms.

Table 2: Covariates balance check, sample of interested in education

	Human mean	Bot mean	Control mean	Human vs Control diff.	Bot vs Control diff.
Female	0.686	0.680	0.684	0.002	-0.004
	(0.464)	(0.467)	(0.465)		
Public high school	0.370	0.373	0.372	-0.002	0.001
	(0.483)	(0.484)	(0.483)		
Private-subsidized high school	0.572	0.573	0.572	0.000	0.001
	(0.495)	(0.495)	(0.495)		
Private high school	0.057	0.055	0.056	0.001	-0.001
	(0.233)	(0.227)	(0.231)		
High school GPA, top 30%	0.333	0.338	0.347	-0.014*	0.009
	(0.471)	(0.473)	(0.476)		
Recent high school graduation	0.776	0.768	0.777	-0.001	-0.009
	(0.417)	(0.422)	(0.416)		
High school GPA	5.799	5.810	5.808	-0.009	0.002
	(0.497)	(0.500)	(0.494)		
High school ranking score	600.8	602.8	602.6	-1.8	0.2
	(126.0)	(126.2)	(124.8)		
Average math and verbal test	486.2	486.4	485.8	0.4	0.6
	(91.4)	(92.6)	(92.7)		
Parents without high school education	0.223	0.222	0.223	0.000	-0.001
	(0.416)	(0.416)	(0.417)		
Parents with high school education	0.595	0.602	0.595	0.000	0.007
	(0.491)	(0.490)	(0.491)		
Parents with higher education	0.182	0.175	0.181	0.001	-0.006
	(0.386)	(0.380)	(0.385)		
Low-income family	0.645	0.654	0.656	-0.011	-0.002
	(0.479)	(0.476)	(0.475)		

Notes: For each covariate, the number of observations is equivalent to as reported in Table 1. Robust tandard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Covariates balance check, sample of interested in social sciences

	Bot mean	Control mean	Bot vs Control diff.
Female	0.639	0.637	0.002
	(0.480)	(0.481)	
Public high school	0.299	0.310	-0.011
	(0.458)	(0.463)	
Private-subsidized high school	0.594	0.576	0.018**
	(0.491)	(0.494)	
Private high school	0.107	0.113	-0.006
	(0.309)	(0.317)	
High school GPA, top 30%	0.334	0.333	0.001
	(0.472)	(0.471)	
Recent high school graduation	0.856	0.863	-0.007
	(0.351)	(0.343)	
High school GPA	5.876	5.879	-0.003
	(0.494)	(0.492)	
High school ranking score	617.5	617.8	-0.3
	(124.4)	(123.2)	
Average math and verbal test	502.0	503.6	-1.6
	(88.6)	(88.2)	
Parents without high school education	0.174	0.173	0.001
	(0.379)	(0.379)	
Parents with high school education	0.577	0.573	0.004
	(0.494)	(0.495)	
Parents with higher education	0.249	0.253	-0.004
	(0.433)	(0.435)	
Low-income family	0.587	0.599	-0.012
	(0.492)	(0.490)	

Notes: For each covariate, the number of observations is equivalent to as reported in Table 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

We rely on three pre-registered outcomes regarding student choices: **Ranked an education major as first option**: takes a one if the student ranked an education major as their first preference of degree to pursue. This outcome is particularly relevant as it indicates a strong preference for pursuing an education major. **Proportion of education majors in the choice set**: this is the proportion of education majors out of the total number of degree options a student selects. This outcome is also relevant because if a student includes more education majors in their choice set, their odds of enrolling as an education major are higher. **Applied to at least one education major**: takes a one if the student applied to at least one education major in their choice set.

We supplement student preferences with an additional variable related to student enrollment in an education major. The outcome, **Enrolled in an education major**, takes a one if the student ultimately enrolled in an education major. However, we interpret this secondary outcome cautiously for several reasons. Firstly, final allocation is a general equilibrium outcome that depends on numerous variables beyond students' preferences and choices, as vacancies are assigned using a deferred acceptance algorithm (Gale and Shapley, 1962). Secondly, the treatment effect on choices could spill over into the allocation patterns of the control group students.

To measure the effectiveness of our interventions, we analyze the students' application and enrollment outcomes, as defined above, for specific academic programs classified by the Higher Education Information Service (SIES) of the Chilean Ministry of Education. SIES maintains a national information system for higher education by collecting information from all institutions and categorizing their academic programs based on areas of knowledge according to ISCED-UNESCO guidelines (UNESCO, 2012). One such area is "education," which includes pedagogy programs (to teach in a classroom) and special education programs (to support students with learning disabilities). Pedagogy programs are further classified into eleven subdisciplines, which we group into three categories: preschool pedagogy, primary education pedagogy, and specialized pedagogy, encompassing the remaining nine pedagogies in specific fields (science, physics, math, history, language arts, foreign languages, arts, philosophy, and physical education).

We present the results computed using the same framework (first-choice, proportion, at least one, enrollment) for specific categories of undergraduate academic programs: education programs (including pedagogy and special education), preschool pedagogy, primary education pedagogy, and specialized pedagogy. The 45 universities participating in the centralized higher education admission system all offer these academic programs.

Table 4 presents descriptive statistics for the outcome variables used in this paper. On average, 12.5 percent of the choice set consisted of applications to education programs. Of the sample, 13 percent applied to education majors as their first choice, while 20.9 percent applied to at least one education major. At the end of the process, 10.6 percent of the sample enrolled in an education degree program.

Table 4: Summary statistics, outcome variables

	Obs.	Mean	Std. dev.	Min	Max
Proportion applied of the choice set					
Education programs	39,119	0.125	0.286	0	1
Preschool pedagogy	39,119	0.022	0.122	0	1
Primary pedagogy	39,119	0.017	0.094	0	1
Specialized pedagogy	39,119	0.029	0.167	0	1
Application as first choice					
Education programs	39,119	0.131	0.337	0	1
Preschool pedagogy	39,119	0.024	0.152	0	1
Primary pedagogy	39,119	0.016	0.127	0	1
Specialized pedagogy	39,119	0.090	0.287	0	1
Application at least once					
Education programs	39,119	0.209	0.406	0	1
Preschool pedagogy	39,119	0.048	0.213	0	1
Primary pedagogy	39,119	0.046	0.210	0	1
Specialized pedagogy	39,119	0.172	0.377	0	1
Enrollment					
Education programs	39,119	0.106	0.308	0	1
Preschool pedagogy	39,119	0.016	0.127	0	1
Primary pedagogy	39,119	0.013	0.115	0	1
Specialized pedagogy	39,119	0.076	0.265	0	1

To assess the overall impact of the intervention on different students' career decisions, our regressions have the following structure:

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

where y_i represents the outcomes (described in section 3.2) of each student i. T_i is a set of dummy variables indicating whether applicant i received the *human* or *bots* treatment, with the control group as a comparison. X_i is a vector including the covariates used for the stratification: a dummy indicating if the student is in the top 30% of her class by GPA, gender, and type of high school. We also control for a dummy variable that takes a value of one if the student was originally interested in education and zero if they were interested in social sciences. Since randomization was conducted at the individual level with no clustering, we report all the results using robust standard errors.

4 Results

Table 5 presents the main results. We document a 1.25 percentage point increase in the probability that a student ranks an education program as their first choice, significant at 5%, in the *humans* arm. That effect represents an increase of 9.5 percent compared to the baseline of 13.1 percent. The *humans* arm also increased the proportion of education programs included in students' choice sets by 1.28 percentage points (also significant at 5%), representing an increase of 10.2 percent compared to the baseline of 12.5 percent, but did not significantly affect the probability of listing at least one education program. This result suggests that the intervention worked mostly on the intensive margin (students seriously considering education in the first place), but not very much on the extensive margin (students not considering education majors to begin with). As a result, the probability that a student in the *humans* arm enrolled in an education program increased by 0.9 percentage points (significant at 10%), 8.5 percent above the baseline of 10.7 percent. In a typical year, 24,000 students are interested in education. Among them, approximately 12% would choose an education major as their first choice without an intervention (2,880). Thus, a 1.25 percentage point increase in the proportion of students ranking education programs as their first choice is equivalent to an increase from 2,880 to 3,150. Considering the caveats related to the enrollment outcome, the treatment effect would imply an increase from 1,848 to 2,069 students enrolled in education majors. The difference represents 0.8% of the projected teacher deficit in 2025. This is a conservative estimate since we are only considering students interested in education as the target population. If we also included students interested in social sciences, the enrollment effect would represent almost 2% of the projected teacher deficit in 2025.

The effects of the *bots* treatment are generally positive but insignificant. The probability that a student ranked an education program as their first choice rose by 0.38 percentage points (insignificant), which represents approximately one-third of the effect of the *humans* arm. A similar pattern emerges when analyzing the effect on the proportion of education programs included in students' choice sets: an insignificant increase of 0.42 percentage points, again about a third of the corresponding *humans* effect. The effect on the probability that students included at least one education program in their choice sets increased by 0.67 percentage points (significant at 10%), a point estimate that is larger than the *humans* arm. The effect on the probability that students enrolled in an education program increased by 0.44 percentage points (insignificant), less than half of the point estimate corresponding to the *humans* equivalent.

Table 5: Preferences and enrollment in education programs

Education programs	Application			
	First choice	Proportion	At least once	Enrollment
Human effect	0.0125**	0.0128**$^{\top}$	0.0099	0.0092*
	(0.0060)	(0.0050)	(0.0068)	(0.0054)
Bot effect	0.0038	0.0042	0.0070*	0.0044
	(0.0034)	(0.0028)	(0.0041)	(0.0031)
Observations	39,005	39,005	39,005	39,005
R-squared	0.088	0.100	0.102	0.077
Dependent variable mean	0.131	0.125	0.209	0.107

Notes: All regressions control for gender, high school characteristics (public, private-subsidized, or private), a dummy indicating if the student was initially interested in education or social sciences, and a top 30% high school GPA dummy. $^{\top}$ means that a point estimate is significant at the 5% level when applying Holm (1979)'s correction for multiple hypotheses. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6 shows the effects by type of education program. The treatments were not intended to promote a specific type of major within education. On the contrary, the idea was to reinforce students' ex-ante motivations and interests. Therefore, we do not have any prior in terms of which major within education could drive the results. With that caveat, preschool education majors consistently explain most of the results. The effects are sizeable. For instance, the *humans* treatment increased the probability of ranking a preschool education program as first choice by almost 30 percent compared to the baseline (2.5%), and the probability of enrolling in that type of program increased by 20 percent concerning the baseline (2%). Although our intervention was not particularly focused on any specific type of major within education, the fact that it seemed to work particularly well among preschool programs is especially encouraging, since evidence indicates that effective teachers have an even greater impact in the earlier years (Heckman et al., 2013; Chetty et al., 2011).[5]

[5] Tables A2 and A3 (Appendix) display equivalent results when the sample is restricted to students who expressed interest in education. Furthermore, Table A4 (Appendix) shows the main results with no controls. In every case, the results are very similar.

Table 6: Preferences and enrollment by type of pedagogy programs

Preschool pedagogy	Application			
	First choice	Proportion	At least once	Enrollment
Human effect	0.0072**ᵀ	0.0045*	0.0065*	0.0040*
	(0.0029)	(0.0023)	(0.0038)	(0.0024)
Bot effect	-0.0008	-0.0007	0.0002	-0.0008
	(0.0015)	(0.0012)	(0.0022)	(0.0013)
Observations	39,005	39,005	39,005	39,005
R-squared	0.025	0.033	0.045	0.020
Dependent variable mean	0.0236	0.0223	0.0479	0.0164

Primary pedagogy	Application			
	First choice	Proportion	At least once	Enrollment
Human effect	-0.0017	0.0017	0.0100*** ᵀ	0.0001
	(0.0023)	(0.0017)	(0.0038)	(0.0021)
Bot effect	0.0000	0.0008	0.0030	0.0006
	(0.0013)	(0.0010)	(0.0021)	(0.0012)
Observations	39,005	39,005	39,005	39,005
R-squared	0.013	0.021	0.030	0.012
Dependent variable mean	0.0165	0.0166	0.0463	0.0135

Specialized pedagogy	Application			
	First choice	Proportion	At least once	Enrollment
Human effect	0.0064	0.0066**	0.0031	0.0044
	(0.0051)	(0.0031)	(0.0063)	(0.0047)
Bot effect	0.0045	0.0016	0.0058	0.0047*
	(0.0029)	(0.0017)	(0.0039)	(0.0027)
Observations	39,005	39,005	39,005	39,005
R-squared	0.061	0.022	0.085	0.054
Dependent variable mean	0.0906	0.0286	0.172	0.0763

Notes: All regressions control for gender, high school characteristics (public, private-subsidized, or private), a dummy indicating if the student was initially interested in education or social sciences, and a top 30% high school GPA dummy. ᵀ means that a point estimate is significant at the 5% level when applying Holm (1979)'s correction for multiple hypotheses. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

As discussed in section 3.2, the interpretation of the outcomes related to students' choices is clear. However, the enrollment results should be interpreted with caution. First, given that the assignment is a general equilibrium variable that depends on a deferred acceptance algorithm, enrollment is affected by factors unrelated to student preferences. Second, there could be spillovers, as student choices in one group (either treatment or control) could affect the others. With this caveat in mind, the results positively affect students' preferences and final enrollment for the *humans* arm and very mild and insignificant results for the *bots* arm.

5 Which students drive the results?

A key policy question is which student groups were impacted by the treatment. Ideally, we would like to see an effect on high-performing students. The literature shows that high-quality teachers are critical for student learning, particularly for disadvantaged or low-income students (Goldhaber, 2002; Rockoff, 2004; Rivkin et al., 2005; Chetty et al., 2014). Therefore, increasing the number of high-quality teachers could significantly impact student achievement and reduce potential inequality. Understanding other factors affecting the recruitment of teachers, such as gender, socioeconomic status, and educational background, could also help improve future campaigns to attract students to education careers.

In our pre-analysis plan, we intentionally omitted any hypotheses regarding heterogeneity. Instead, we planned to utilize the size and richness of our dataset to leverage causal forest estimators to identify heterogeneous effects (Athey and Imbens, 2017). This approach reduces researcher discretion in selecting relevant dimensions of heterogeneity, allowing the data to speak for itself. Following the "honest" method developed by Athey et al. (2019), we estimate Conditional Average Treatment Effects (CATE) for each individual in our sample using a generalized random forest (grf R package). By relying on data-driven sample splits, this method limits researcher discretion when selecting the relevant dimensions of heterogeneity. We estimate CATE for each individual based on all characteristics included in our data.

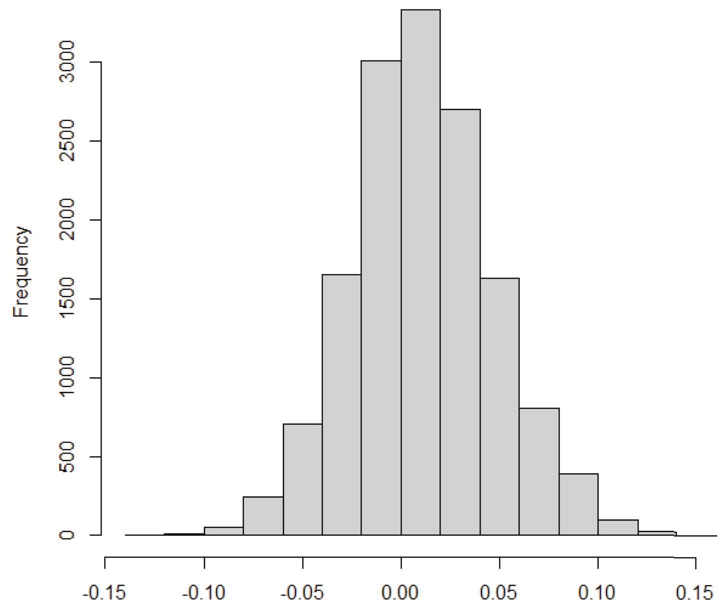
Causal forest algorithms are adaptations of regression trees commonly used in the machine learning literature (Breiman et al., 2017). Regression trees are recursive partitioning algorithms that split a sample to maximize heterogeneity across splits. A forest is a group of trees, and each tree is grown from a portion of the data drawn randomly from the full sample.

For each tree in our experiment, we select a sub-sample randomly without replacement from the full sample. The tree is grown by recursively splitting each node into child nodes, with splits chosen by the algorithm to maximize the heterogeneity of the average treatment effect in each subgroup. A node becomes a final leaf when a split no longer improves the fit. To avoid overfitting, the honest approach, recommended by Athey and Imbens (2016), splits the randomly selected sub-sample into two halves: one used to grow the tree and the other to estimate the average treatment effect within each leaf. This helps reduce overestimation of the model's goodness of fit, naturally reducing statistical power.

Our analysis focuses on this group since we only found positive results in the *humans* arm. We examine the effects on three primary outcomes related to student preferences: (1) application to an education program as first choice, (2) proportion of education majors included in choice set, and (3) application to at least one education program. We exclude the enrollment outcome due to the potential interpretation issues discussed earlier.

Figure 1 displays the predicted treatment effects distribution using the **application to an education program as first choice** outcome. The median point estimate is approximately 1.5%, which is close to the average treatment effect of our main estimates, but with significant heterogeneity. The algorithm identified several covariates that frequently appear in the splits. Student performance variables, such as the student's ranking (appears in 37% of splits), GPA (17%), and university entrance exam scores (17%), are the most relevant. Gender appears in 5% of the splits, the low-income dummy appears in 2% of the splits, and the other covariates appear in less than 2% of the splits.

Figure 1: Histogram of conditional average treatment effects (CATE)



Notes: This histogram displays the distribution of CATE across the entire out-of-bag sample using the outcome **application to an education program as first choice**.

To analyze the different characteristics that maximize the effects, we estimate the treatment effects for the 20% of students for which the treatment was least effective (Q1) and the 20% of students for which the treatment was most effective (Q5). Table 7 presents the mean of each covariate for Q1 and Q5. In each table, we report the sample mean, the mean of the covariate for the individuals comprising each group (least or most affected by the treatment), and the standardized difference between the means of each group. A few interesting patterns emerge.[6]

Across all outcomes related to student choices, we consistently observe a higher proportion of male students in the group more affected by the treatment. For example, while the proportion of female students in Q1 for the outcome **application to an education program as first choice** was 71.1%, it was only 62.4% in Q5. This finding is noteworthy considering the gender imbalances in the teaching profession (OECD, 2017; Elacqua et al., 2022b). We also observe a slightly higher proportion of students from private schools in Q1 compared to Q5 for all outcomes, indicating that the treatment was more effective among students from non-private high schools (public and private-

[6]Considering that inference could be problematic for testing the difference between the means of the two groups, we follow Britto et al. (2022) and analyze the magnitude of the standardized difference in each case qualitatively.

subsidized), which is consistent with family income. Although the proportion of low-income students is slightly higher in Q5 versus Q1 for every outcome, this variable is based on self-reported family income and includes a large proportion of missing values. These results are consistent with previous research (Ajzenman et al., 2022), and reflect the perception that education is a more prestigious profession among low-income families in Chile.

Finally, the three variables related to student ability point in the same direction. High school ranking, which measures a student's relative performance within their high school cohort, high school GPA, and math and verbal test scores on the university entrance exams are notably higher in Q5 than Q1. An important caveat is that the first two variables are much less informative because they are not standard (different schools may assess, grade, and rank students differently). The last outcome, however, is standardized and comparable. The average scores on the mandatory components of the university entrance exam are 504 in Q5 and 481 (almost 5% lower) in Q1 for the outcome **application to an education program as first choice**. This holds for every outcome related to student choices: higher-performing students tend to be those for whom the treatment was particularly effective. This finding has significant policy implications, as better-qualified teachers (high-performing candidates) will likely be more effective at improving student learning (Chilean evidence shows that higher admission scores are good predictors of higher value-added, see Neilson et al. (2019)).

The results of the three panels in Table 7 suggest that a few covariates play a crucial role in explaining the observed heterogeneity. Indeed, only three variables exhibit standardized differences larger than 0.2 SD (a typical critical value, see Britto et al. (2022)): the three covariates related to student performance. Notably, gender appears to be borderline relevant, registering above 0.2 in one case and ranging between 0.1 and 0.2 in others.

Table 7: Predicted conditional average treatment effects (CATE) of the human treatment in education programs

Application as first choice	(1) Sample mean	(2) CATE Q1	(3) CATE Q5	(4) Std. diff. (2)-(3)
Female	0.687	0.711	0.624	0.188
Public high school	0.367	0.343	0.349	-0.013
Private-subsidized high school	0.574	0.595	0.603	-0.016
Private high school	0.058	0.062	0.048	0.060
High school ranking score	604.2	581.9	638.3	-0.451
Recent high school graduation	0.782	0.750	0.793	-0.102
High school GPA	5.814	5.736	5.942	-0.419
Average math and verbal	488.2	481.4	504.8	-0.272
Parents without high education	0.219	0.156	0.180	-0.057
Parents with high education	0.598	0.684	0.627	0.117
Parents with higher education	0.183	0.160	0.194	-0.087
Low-income family	0.571	0.548	0.614	-0.133

Proportion of the choice set	(1) Sample mean	(2) CATE Q1	(3) CATE Q5	(4) Std. diff. (2)-(3)
Female	0.687	0.738	0.596	0.306
Public high school	0.367	0.372	0.327	0.092
Private-subsidized high school	0.574	0.561	0.616	-0.111
Private high school	0.058	0.067	0.057	0.044
High school ranking score	604.2	576.7	632.0	-0.442
Recent high school graduation	0.782	0.729	0.783	-0.132
High school GPA	5.814	5.710	5.924	-0.433
Average math and verbal	488.2	477.0	505.5	-0.332
Parents without high education	0.219	0.136	0.177	-0.100
Parents with high education	0.598	0.678	0.641	0.074
Parents with higher education	0.183	0.186	0.181	0.013
Low-income family	0.571	0.565	0.620	-0.111

Application at least once	(1) Sample mean	(2) CATE Q1	(3) CATE Q5	(4) Std. diff. (2)-(3)
Female	0.687	0.703	0.642	0.131
Public high school	0.367	0.384	0.316	0.140
Private-subsidized high school	0.574	0.544	0.634	-0.182
Private high school	0.058	0.073	0.050	0.097
High school ranking score	604.2	591.5	625.6	-0.273
Recent high school graduation	0.782	0.764	0.801	-0.090
High school GPA	5.814	5.765	5.900	-0.274
Average math and verbal	488.2	475.3	504.3	-0.337
Parents without high education	0.219	0.145	0.239	-0.230
Parents with high education	0.598	0.669	0.608	0.126
Parents with higher education	0.183	0.186	0.153	0.086
Low-income family	0.571	0.511	0.636	-0.253

Notes: Column (2) represents the average effect of each covariate for observations in the bottom 20% (Q1) of treatment impact. Column (3) shows the average effect of each covariate for observations in the top 20% (Q5) of treatment impact. Column (4) shows a difference of means test between columns (2) and (3), normalized in terms of standard deviations.

6 Discussion

Effective teachers are a crucial component of the education production process (Rivkin et al., 2005; Kane and Staiger, 2008; Bau and Das, 2020). However, shortages in the supply of high-quality teachers have become an increasingly challenging issue in many countries (Elacqua et al., 2022b; Bertoni et al., 2018), which can compromise the quality of education and potentially have lasting impacts on key development outcomes (Chetty et al., 2014; Araujo et al., 2016).

Several countries, including Chile, have tried to improve the objective conditions of education-related careers (Ávalos and Bellei, 2019; Pérez-Núnez, 2020). However, teacher shortages remain a persistent problem, and low-cost interventions based on insights from behavioral economics targeting specific career choice factors could complement structural reforms. Our study shows that human-intensive tutoring campaigns can effectively promote education careers. The main results of the experiment are generally positive, as the *humans* intervention is still a cost-effective approach, particularly when compared to other policies such as tuition scholarships or improving salaries and working conditions. The human-intensive intervention had positive and substantial effects, which could lead to a reduction of between 0.8 and 2 percent of the projected 2025 Chilean teacher deficit if scaled up. It was particularly effective among high-performing students, a critical factor in improving the quality of teacher supply, and male students, which is an important finding given the significant gender imbalance in the teaching profession (Elacqua et al., 2022b). Our results show that a light-touch intervention can successfully increase the quality and quantity of prospective teachers, particularly relevant in the context of significant teacher shortages.

Still, scalability remains a challenge. While the human-intensive intervention had positive and sizeable effects, a scalable equivalent failed to produce meaningful change. Unfortunately, there could be many explanations for this failure, and isolating any specific factor is nearly impossible.

One plausible hypothesis for the failure of the scalable chatbot program is that its implementation was ineffective. This does not seem to be the case in our setting. If any-thing, the proportion of students successfully contacted by Chatbots was slightly higher (93.6%) than the equivalent in the *humans* arm (92%). Another possibility is that, after the first successful contact, *humans* were more effective in providing personalized and nuanced responses to students' inquiries, whereas the chatbot responses may have been too generic or inflexible to address student concerns fully.

However, our analysis suggests this was not the case, although comparing the two treatments is difficult as they were different in nature. With that said, only 32% of conversations initiated by human tutors were successfully completed (that is, 68% did not finish the planned script), while in the case of Chatbots, approximately 14% of messages were rejected by students, and around 33% of conversations ended because the messages were unanswered three consecutive times (which is the equivalent of not being able to finish a complete script), totaling 47%. Our data do not allow us to analyze the moment a conversation ended in the *humans* arm, so it is impossible to rule out the possibility that human conversations (and exposure) were more extended. Nevertheless, the significantly lower proportion of unfinished scripts in the case of Chatbots (47% versus 68%) makes this hypothesis less likely.

If both treatments were implemented relatively successfully, an alternative explanation could be that Chatbots were not as persuasive as humans. However, this does not imply that Chatbots are generally ineffective or that every type of chatbot will be ineffective in our context. The fact that the specific Chatbots we used, which are particularly low-cost and easy to implement, did not work does not necessarily mean that other types of *bots* will not be effective.

Given how successful the *humans* intervention was (a large effect at a relatively low cost, compared to more structural interventions), a promising avenue is to explore ways to improve the program's scalability. For instance, one way of enhancing the impact of Chatbots is to explore AI-trained *bots*, which have proven to be effective in other contexts (Page and Gehlbach, 2017; Luo et al., 2019; Nurshatayeva et al., 2021) and might create interactions that are more similar to human interactions. While they may be costlier and more challenging to implement than rule-based *bots*, they could represent a plausible way to scale up a program that, when delivered by *humans*, significantly impacts students' choices.

References

Abd-Alrazaq, A. A., Rababeh, A., Alajlani, M., Bewick, B. M. and Househ, M. (2020). Effectiveness and safety of using chatbots to improve mental health: Systematic review and meta-analysis, *Journal of Medical Internet Research* **22**(7).

Ajzenman, N., Bertoni, E., Elacqua, G., Marotta, L. and Méndez, C. (2023). Altruism or money? Reducing teacher sorting using behavioral strategies in Peru, *Journal of Labor Economics* .

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., Marotta, L. and Olsen, A. (2022). Order effects and employment decisions: Experimental evidence from a nationwide program.

Al-Ubaydli, O., Lee, M. S., List, J. A., Mackevicius, C. L. and Suskind, D. (2021). How can experiments play a greater role in public policy? Twelve proposals from an economic model of scaling, *Behavioural Public Policy* **5**(1): 2–49.

Al-Ubaydli, O., List, J. A., LoRe, D. and Suskind, D. (2017). Scaling for economists: Lessons from the non-adherence problem in the medical literature, *Journal of Economic Perspectives* **31**(4): 125–44.

Al-Ubaydli, O., List, J. A. and Suskind, D. (2020). 2017 Klein lecture: The science of using science: Toward an understanding of the threats to scalability, *International Economic Review* **61**(4): 1387–1409.

Al-Ubaydli, O., List, J. A. and Suskind, D. L. (2017). What can we learn from experiments? Understanding the threats to the scalability of experimental results, *American Economic Review* **107**(5): 282–86.

Angrist, J. D. and Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics, *Journal of Economic Perspectives* **24**(2): 3–30.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y. and Schady, N. (2016). Teacher quality and learning outcomes in kindergarten, *The Quarterly Journal of Economics* **131**(3): 1415–1453.

Athey, S. and Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects, *Proceedings of the National Academy of Sciences* **113**(27): 7353–7360.

Athey, S. and Imbens, G. (2017). The state of applied econometrics: Causality and policy evaluation, *Journal of Economic Perspectives* **31**(2): 3–32.

Athey, S., Tibshirani, J. and Wager, S. (2019). Generalized random forests, *The Annals of Statistics* **47**(2): 1148–1178.

Ávalos, B. and Bellei, C. (2019). Recent education reforms in Chile: How much of a departure from market and new public management systems?, *Politics of education in Latin America*, Brill, pp. 43–71.

Bacolod, M. P. (2007). Do alternative opportunities matter? the role of female labor markets in the decline of teacher quality, *The Review of Economics and Statistics* **89**(4): 737–751.

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., Shotland, M. and Walton, M. (2017). From proof of concept to scalable policies: Challenges and solutions, with an application, *Journal of Economic Perspectives* **31**(4): 73–102.

Baron, J. (2021). Union reform, performance pay, and new teacher supply: Evidence from Wisconsin's Act 10, *AEA Papers and Proceedings*, Vol. 111, pp. 445–49.

Bau, N. and Das, J. (2020). Teacher value added in a low-income country, *American Economic Journal: Economic Policy* **12**(1): 62–96.

Bertoni, E., Elacqua, G., Jaimovich, A., Rodriguez, J. and Santos, H. (2018). Teacher policies, incentives, and labor markets in Chile, Colombia, and Peru: Implications for equality, *Technical Note No IDB-WP-00945. BID* .

Bertoni, E., Elacqua, G., Marotta, L., Martínez, M., Méndez, C., Montalva, V., Westh Olsen, A., Soares, S. and Santos, H. (2020). El problema de la escasez de docentes en Latinoamérica y las políticas para enfrentarlo, *Nota Técnica No IDB-tn-01883. BID* .

Boyd, D., Grossman, P., Lankford, H., Loeb, S. and Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement.

Boyd, D., Lankford, H., Loeb, S. and Wyckoff, J. (2013). Analyzing the determinants of the matching of public school teachers to jobs: Disentangling the preferences of teachers and employers, *Journal of Labor Economics* **31**(1): 83–117.

Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J. (2017). *Classification and regression trees*, Routledge.

Britto, D. G., Pinotti, P. and Sampaio, B. (2022). The effect of job loss and unemployment insurance on crime in Brazil, *Econometrica* **90**(4): 1393–1423.

Castro-Zarzur, R., Espinoza, R. and Sarzosa, M. (2022). Unintended consequences of free college: Self-selection into the teaching profession, *Economics of Education Review* **89**: 102260.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W. and Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR, *The Quarterly Journal of Economics* **126**(4): 1593–1660.

Chetty, R., Friedman, J. N. and Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood, *American Economic Review* **104**(9): 2633–2679.

Clotfelter, C., Glennie, E., Ladd, H. and Vigdor, J. (2008a). Teacher bonuses and teacher retention in low-performing schools: Evidence from the North Carolina $ 1,800 teacher bonus program, *Public Finance Review* **36**(1): 63–87.

Clotfelter, C., Glennie, E., Ladd, H. and Vigdor, J. (2008b). Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina, *Journal of Public Economics* **92**(5-6): 1352–1370.

Corcoran, S. P., Evans, W. N. and Schwab, R. M. (2004). Changing labor-market opportunities for women and the quality of teachers, 1957-2000, *American Economic Review* **94**(2): 230–235.

De Ree, J., Muralidharan, K., Pradhan, M. and Rogers, H. (2018). Double for nothing? Experimental evidence on an unconditional teacher salary increase in indonesia, *The Quarterly Journal of Economics* **133**(2): 993–1039.

DellaVigna, S. and Linos, E. (2022). RCTs to scale: Comprehensive evidence from two nudge units, *Econometrica* **90**(1): 81–116.

Elacqua, G., Hincapie, D., Hincapié, I. and Montalva, V. (2022a). Can financial incentives help disadvantaged schools to attract and retain high-performing teachers? Evidence from Chile, *Journal of Policy Analysis and Management* **41**(2): 603–631.

Elacqua, G., Jaimovich, A., Pérez-Núnez, G., Hincapié, D., Gómez, C., Sánchez, M., Escalona, G. and Walker, J. (2022b). Who studies pedagogy? Trends in the profile of future teachers in Latin America and the Caribbean, *Technical Note No IDB-TN-02619. BID* .

Falch, T. (2011). Teacher mobility responses to wage changes: Evidence from a quasi-natural experiment, *American Economic Review* **101**(3): 460–65.

Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage, *The American Mathematical Monthly* **69**(1): 9–15.

Goldhaber, D. (2002). The mystery of good teaching, *Education next* **2**(1): 50–55.

Han, S. W., Borgonovi, F. and Guerriero, S. (2018). What motivates high school students to want to be teachers? The role of salary, working conditions, and societal evaluations about occupations in a comparative perspective, *American Educational Research Journal* **55**(1): 3–39.

Heckman, J., Pinto, R. and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes, *American Economic Review* **103**(6): 2052–2086.

Hendricks, M. D. (2014). Does it pay to pay teachers more? Evidence from Texas, *Journal of Public Economics* **109**: 50–63.

Holm, S. (1979). A simple sequentially rejective multiple test procedure, *Scandinavian Journal of Statistics* pp. 65–70.

Hoxby, C. M. and Leigh, A. (2004). Pulled away or pushed out? Explaining the decline of teacher aptitude in the United States, *American Economic Review* **94**(2): 236–240.

Hoyle, E. (2001). Teaching: Prestige, status and esteem, *Educational management & administration* **29**(2): 139–152.

Jackson, C. K. (2009). Student demographics, teacher sorting, and teacher quality: Evidence from the end of school desegregation, *Journal of Labor Economics* **27**(2): 213–256.

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non–test score outcomes, *Journal of Political Economy* **126**(5): 2072–2107.

Kane, T. J. and Staiger, D. O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation, *Technical report*, National Bureau of Economic Research.

Lankford, H., Loeb, S. and Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis, *Educational Evaluation and Policy Analysis* **24**(1): 37–62.

List, J. A. (2022). *The voltage effect: How to make good ideas great and great ideas scale*, Currency.

Luo, X., Tong, S., Fang, Z. and Qu, Z. (2019). Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases, *Marketing*

Science **38**(6): 937–947.

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.

Muralidharan, K. and Niehaus, P. (2017). Experimentation at scale, *Journal of Economic Perspectives* **31**(4): 103–24.

Neilson, C., Gallegos, S., Calle, F. and Karnani, M. (2019). Screening and recruiting talent at teacher colleges using pre-college academic achievement, *Technical report*.

Nurshatayeva, A., Page, L. C., White, C. C. and Gehlbach, H. (2021). Are artificially intelligent conversational chatbots uniformly effective in reducing summer melt? Evidence from a randomized controlled trial, *Research in higher education* **62**: 392–402.

OECD (2017). Gender imbalances in the teaching profession, *Education Indicators in Focus* (49).

Page, L. C. and Gehlbach, H. (2017). How an artificially intelligent virtual assistant helps students navigate the road to college, *Aera Open* **3**(4): 2332858417749220.

Pérez-Núnez, G. (2020). Building a teacher pipeline: Evidence from a merit-based tuition scholarship, *Advancing Equity: The Effects of Higher Education Interventions for Improving Underserved Students' Educational Outcomes*, Stanford University, pp. 112–158.

Perryman, J. and Calvert, G. (2020). What motivates people to teach, and why do they leave? Accountability, performativity and teacher retention, *British Journal of Educational Studies* **68**(1): 3–23.

Rivkin, S. G., Hanushek, E. A. and Kain, J. F. (2005). Teachers, schools, and academic achievement, *Econometrica* **73**(2): 417–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data, *American Economic Review* **94**(2): 247–252.

Skatova, A. and Ferguson, E. (2014). Why do different people choose different university degrees? Motivation and the choice of degree, *Frontiers in Psychology* **5**: 1244.

Staiger, D. O. and Rockoff, J. E. (2010). Searching for effective teachers with imperfect information, *Journal of Economic Perspectives* **24**(3): 97–118.

UNESCO (2012). International standard classification of education: Isced 2011, *Comparative Social Research* **30**.

Vivalt, E. (2020). How much can we generalize from impact evaluations?, *Journal of the*

European Economic Association **18**(6): 3045–3089.

Watt, H. M., Richardson, P. W., Klusmann, U., Kunter, M., Beyer, B., Trautwein, U. and Baumert, J. (2012). Motivations for choosing teaching as a career: An international comparison using the fit-choice scale, *Teaching and teacher education* **28**(6): 791–805.

Appendix A:

Table A1: Sample stratification

Groups	Size	Gender		High School type			High School Performance
		Male	Female	Public	Priv. Sub.	Private	Top 30%
Full sample	40,692	13,627	27,065	13,941	23,147	3,078	13,640
		33%	67%	35%	58%	8%	34%
Interested in education	24,429	7,740	16,689	9,022	13,755	1,337	8,227
		32%	68%	37%	57%	6%	34%
Human	8,135	2,555	5,580	3,003	4,589	449	2,685
		31%	69%	37%	57%	6%	33%
Bot	8,148	2,609	5,539	3,030	4,575	435	2,719
		32%	68%	38%	57%	5%	33%
Control	8,146	2,576	5,570	2,989	4,591	453	2,823
		32%	68%	37%	57%	6%	35%
Interested in social sciences	16,263	5,887	10,376	4,919	9,392	1,741	5,413
		36%	64%	30%	58%	11%	33%
Bot	8,132	2,938	5,194	2,429	4,768	833	2,705
		36%	64%	30%	59%	10%	33%
Control	8,131	2,949	5,182	2,490	4,624	908	2,708
		36%	64%	31%	58%	11%	33%

Table A2: Preferences and Enrollment: Education Majors (only interested in education)

	Education programs			
	Rankfirst	Proportion	Listed at least once	Enrolled
Human effect	0.0135**	0.0142*** ⊤	0.0121	0.0121**
	(0.0065)	(0.0055)	(0.0074)	(0.0059)
Bot effect	0.0053	0.0066	0.0103	0.0093
	(0.0065)	(0.0054)	(0.0073)	(0.0059)
Observations	23,231	23,231	23,231	23,231
R-squared	0.007	0.011	0.015	0.026
Mean of dependent variable	0.212	0.197	0.312	0.170

Notes: All regressions control for gender, high school characteristics (public, private-subsidized, or private), and a top 30% high school GPA dummy. ⊤ means that a point estimate is significant at the 5% level when applying Holm (1979)'s correction for multiple hypotheses. Education programs follow CINE-UNESCO categories (UNESCO, 2012), including psychopedagogy and education programs. Education programs are split into eleven specialties. Related degrees consist of bachelor's degrees that could lead to an education degree. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Preferences and Enrollment: Education Majors by type (only interested in education)

Panel A	Preschool pedagogy			
	Rankfirst	Proportion	Listed at least once	Enrolled
Human effect	0.0066**^	0.0038^	0.0066	0.0038^
	(0.0031)	(0.0025)	(0.0042)	(0.0026)
Bot effect	-0.0019	-0.0020	0.0007	-0.0014
	(0.0029)	(0.0024)	(0.0041)	(0.0024)
Observations	23,231	23,231	23,231	23,231
R-squared	0.018	0.025	0.035	0.018
Mean of dependent variable	0.0384	0.0358	0.0743	0.0261

Panel B	Primary Pedagogy			
	Rankfirst	Proportion	Listed at least once	Enrolled
Human effect	-0.0016	0.0023	0.0119*** $^{\top}$	0.0007
	(0.0026)	(0.0019)	(0.0041)	(0.0023)
Bot effect	0.0003	0.0020	0.0067*	0.0017
	(0.0026)	(0.0019)	(0.0040)	(0.0023)
Observations	23,231	23,231	23,231	23,231
R-squared	0.005	0.008	0.011	0.007
Mean of dependent variable	0.0270	0.0266	0.0724	0.0220

Panel C	Specialized pedagogy			
	Rankfirst	Proportion	Listed at least once	Enrolled
Human effect	0.0079	0.0076**	0.0048	0.0071
	(0.0056)	(0.0034)	(0.0069)	(0.0052)
Bot effect	0.0068	0.0035	0.0078	0.0092*
	(0.0056)	(0.0033)	(0.0069)	(0.0052)
Observations	23,231	23,231	23,231	23,231
R-squared	0.012	0.006	0.021	0.020
Mean of dependent variable	0.146	0.0472	0.256	0.122

Notes: All regressions control for gender, high school characteristics (public, private-subsidized, or private), and a top 30% high school GPA dummy. $^{\top}$ means that a point estimate is significant at the 5% level when applying Holm (1979)'s correction for multiple hypotheses. Education programs follow CINE-UNESCO categories (UNESCO, 2012), including psychopedagogy and education programs. Education programs are split into eleven specialties. Related degrees consist of bachelor's degrees that could lead to an education degree. Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1.

Table A4: Preferences and Enrollment: Education Majors (no controls)

	Education programs			
	Rankfirst	Proportion	Listed at least once	Enrolled
Human effect	0.0118**	0.0121** $^{\top}$	0.0086	0.0080
	(0.0060)	(0.0051)	(0.0068)	(0.0055)
Bot effect	0.0038	0.0041	0.0069*	0.0041
	(0.0033)	(0.0028)	(0.0041)	(0.0031)
Observations	39,119	39,119	39,119	39,119
R-squared	0.084	0.094	0.093	0.063
Mean of dependent variable	0.131	0.125	0.209	0.106

Notes: Regressions do not include any controls. $^{\top}$ means that a point estimate is significant at the 5% level when applying Holm (1979)'s correction for multiple hypotheses. Education programs follow CINE-UNESCO categories (UNESCO, 2012), including psychopedagogy and education programs. Education programs are split into eleven specialties. Related degrees consist of bachelor's degrees that could lead to an education degree. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix B:

DEMRE survey questions to determine students' eligibility (translated from Spanish

In this section, you will be asked about your reasons for registering for the PDT, how you have prepared for it, the elective courses you took in the last year of high school, and whether you want Elige Educar to have access to your data to guide your vocational decision, among other things.

- **In which area would you like to study or work?**
  - Humanities (philosophy, translation and interpretation, degree in literature, library science, etc.)
  - Art and architecture (design, music, dance, landscaping, digital animation, theater, etc.)
  - Health (medicine, nursing, kinesiology, nutrition, obstetrics, etc.)
  - **Education (teaching, educational psychology, degree in education, educational technician, etc.)**
  - Technology (civil engineering, industrial engineering, industrial mechanics engineering, surveying, industrial design, industrial chemistry, environmental engineering, etc.)
  - **Social sciences (public administration, archaeology, anthropology, journalism, history, psychology, social work, advertising, political science, sociology, etc.)**
  - Agriculture (agronomy, veterinary, forestry, aquaculture and fisheries technician, etc.)
  - Basic sciences (chemical analyst, geology, physics, astronomy, mathematics, etc.)
  - Business administration and commerce (commercial engineering, auditor, logistics engineering, accounting, etc.)
  - Law (law, legal technician, etc.)
  - Not sure
  - Prefer not to answer

- **Do you agree to let us share your contact information with Elige Educar* so they can guide your vocational decision? (full name, ID number, school RBD, graduation year, email address, and telephone number)**
- **Do you agree to be contacted by Elige Educar* to participate in education studies and research?**

  *Note: Elige Educar is a non-profit project that seeks to attract young people to teaching careers. For more information, visit the website: www.eligeeducar.cl

Appendix C:

Summary of conversation flow (bot and *humans*)[7]

MODULE 1: WELCOME AND INITIAL GREETING

Hi, I hope you're doing well. Welcome to the Elige Educar chatbot. A few months ago, through DEMRE, you expressed interest in receiving vocational guidance, and that will be my mission! Would you like us to talk here so I can help you with information for your career choice?

- **No response**: (See no-response protocol)[8]
- **No**: No problem! If you want to receive information at another time, write to us in this same chat and we will be in touch. Take care! :)
- **Yes**: Excellent! To better guide you, I will ask you some questions. Have you already decided on the career you will apply for?

    1. **I already know the career I want to study for**

       Good! Is the career any kind of education degree?

       – **Yes**: [*Message A*] That's great news! Congratulations on being interested in one of the most important professions for society. Being a teacher allows you to accompany thousands in their education, teach the subjects you love the most, build opportunities and shape the future of the country.
         Would you like me to send you information about higher education and pedagogy careers here?

         * **Yes**: *Step 2A motivation*
         * **No**: No problem! If you want to receive information at another time, write to us in this same chat and we will be in touch. Take care! :)
         * **No response**: See no-response protocol

       – **No**: [*Message B*] Would you like me to send you information about the application process and careers in higher education here?

         * **Yes**: *Step 2B motivation*
         * **No**: No problem! If you want to receive information at another time, write to us in this same chat and we will be in touch. Take care! :)
         * **No response**: See no-response protocol

       – **No response**: See no-response protocol

    2. **I am considering two or more options**

       Good! And among your options, have you considered any kind of education degree?

       – **Yes**: [*Message A*]
         Would you like me to send you information about higher education and pedagogy careers here?

[7] Messages were sent weekly to random groups of students, similarly to the pace of human tutors.
[8] The **no response protocol** can be summarized as follows. The maximum number of contact attempts is three times over a period of three weeks. If the person does not respond after the first attempt, they should be contacted for a second time the following week with a greeting. If the person accepts, then the conversation flow is resumed; if they do not accept, they are thanked and the conversation is ended; and if they do not respond, a summary with the missing information from the conversation flow is sent to them the next day.

    * **Yes**: *Step 2A motivation*
    * **No**: No problem! If you want to receive information at another time, write to us in this same chat and we will be in touch. Take care! :)
    * **No response**: See no-response protocol
  - **No**: [*Message B*] Would you like me to send you information about the application process and careers in higher education here?
    * **Yes**: *Step 2B motivation*
    * **No**: No problem! If you want to receive information at another time, write to us in this same chat and we will be in touch. Take care! :)
    * **No response**: See no-response protocol
  - **No response**: See no-response protocol

3. **I haven't decided what to study yet**

   [*Message B*] And would you like me to send you information through here about the application process and higher education careers?
   - **Yes**: *Step 2B motivation*
   - **No**: No problem! If you want to receive information at another time, write to us in this same chat and we will be in touch. Take care! :)
   - **No response**: See no-response protocol

4. **I don't know if I will apply to higher education**

   [*Message B*] And would you like me to send you information through here about the application process and higher education careers?
   - **Yes**: *Step 2B motivation*
   - **No**: No problem! If you want to receive information at another time, write to us in this same chat and we will be in touch. Take care! :)
   - **No response**: See no-response protocol

5. **No response**

   See no-response protocol

MODULE 2: MOTIVATION

- *Step 2A motivation*

  Excellent :) Let's begin! First, I would like to know, what motivates you the most to consider a career in education?

  1. **To contribute to the education of future generations and improve opportunities for children. It is a service to society.**

     Pedagogy will definitely allow you to be an agent of change, working daily to reduce the educational and social gaps in the country. Throughout your professional trajectory, a teacher impacts the lives and opportunities of 5,000 students.

  2. **I enjoy teaching and sharing my knowledge with others.**

     Pedagogy is a great option if you have a particular area of interest and want to share it with others, especially if you have skills for teaching and working with children and teenagers. You will be part of a profession that adapts every day to the challenges of a changing world. It is a great opportunity to have a dynamic and innovative job!

3. **Teaching provides stable employment and/or income, and even more vacation time (due to the school calendar)**.

   Did you know that in teaching you can start earning a salary of almost $1,000,000 (full-time), which increases with experience and good performance in teaching evaluations? The average employability of teachers is similar to careers such as public administration, law, and commercial engineering. You won't have trouble finding work!

4. **Teaching is a prestigious and demanding profession. Only for the best.**

   Did you know that teaching is among the four most highly valued professions in Chile? (along with medicine, engineering, and law). Also, did you know that only teaching, medicine, and dentistry require university accreditation? Teaching is one of the most prestigious careers in society!

5. **No response**: See non-response protocol.

What is the second thing that motivates you the most to consider a career in Education? [Show only menu options not previously selected]

1. To contribute to the education of future generations and improve opportunities for children. It is a service to society.
2. I enjoy teaching and sharing my knowledge with others.
3. Teaching provides stable employment and/or income, and even more vacation time (due to the school calendar).
4. Teaching is a prestigious and demanding profession. Only for the best.
5. None of the above, let's move on to the next question.

- *Step 2B motivation*

  Great :) Let's start! First, I would like to know, what characteristics are you looking for in a career?

1. **That it allows me to help people and make a contribution to society.**

   Teaching is an excellent alternative if what you are looking for is a job with impact. Did you know that in their professional trajectory, a teacher impacts the lives and opportunities of 5,000 students? Through pedagogy, you will be an agent of change, working daily to reduce the educational and social gaps in the country.

2. **That I like it, that it reflects my vocation, and that it allows me to learn about the topics that interest me the most.**

   Pedagogy is a great option if you like a particular area and want to share your interest with others. Especially if you have skills to teach and work with children and adolescents. Those who like science and mathematics share their knowledge in laboratories and school classrooms. In humanities and social sciences, we find excellent teachers of language, philosophy, and history. There are also artists who teach dance, theater, music, and visual arts.

3. **That it allows me to obtain a high salary and good working conditions.**.

   Did you know that in pedagogy, you can start earning a salary of almost $1,000,000 (full-time), which increases with experience and good performance in teaching evaluations? The average employability of pedagogies is similar to careers such as public administration, law, and commercial engineering. You won't have trouble finding work!

4. **A career that is selective and prestigious, and allows me to have a valued position in society.**

   Did you know that pedagogy is among the four most highly valued professions in Chile? (along with medicine, engineering, and law). Also, did you know that only pedagogy, medicine, and dentistry require universities to be accreditation? Pedagogy is one of the most prestigious careers in society!

5. **No response**: See non-response protocol.

What other characteristic are you looking for in a professional career?

[Show only menu options not previously selected]

1. That it allows me to help people and make a contribution to society.
2. That I like it, that it reflects my vocation, and that it allows me to learn about the topics that interest me the most.
3. That it allows me to obtain a high salary and good working conditions.
4. A career that is selective and prestigious, and allows me to have a valued position in society.
5. None of the above, let's move on to the next question.

MODULE 3: QUESTIONS AND CONCERNS REGARDING TEACHER EDUCATION PROGRAMS

Next, I present you with a list of questions you may have about the teaching profession. Please indicate which one you would like more information on:

1. **Will I have what it takes to be a good teacher?**

   The skills needed to be a teacher are developed in university and through practice, so don't worry! The most important thing is that you are interested in working with people, enjoy teaching, and want to have an impactful job. Additionally, in your first years of work, you can rely on the mentorship of an expert teacher who can guide you through your first work experiences.

2. **How do I convince those around me to support me in studying education?**

   The education degree is among the four most valued in Chile, along with medicine, engineering, and law. Since 2017, with Law 20.903, the requirements to become a teacher have increased, making it a more selective and demanding career. Furthermore, the same law has greatly improved working conditions (salary, workload, training opportunities, etc.).

3. **What is the job outlook and employment opportunities for education majors?**

   According to data from www.mifuturo.cl, the average employment rate for education majors is 77% in the first year. This is similar to careers such as Public Administration, Law, and Commercial Engineering. Even Special Education and Mathematics Education have employment rates over 90%. Additionally, research projects that by 2025, Chile will lack more than 26,000 qualified teachers and 6,700 early childhood educators in the country. Chile needs more and better teachers!

4. **Is it true that teachers' salaries are very low?**

   No. In education, you start with a salary of almost $1,000,000 (full-time), which increases with experience and good performance in teaching evaluations. An experienced teacher can earn around 3 million.

5. **Is teaching a monotonous profession?**

   The teaching profession is very challenging and dynamic because each context and each student has their own characteristics. Furthermore, the rapid changes in the world and the new skills of the 21st century demand that teachers constantly innovate in the way they teach. Teachers also teach at universities, conduct research, lead education policies in the Ministry of Education, and work for foundations and organizations that support schools. There are many possibilities!

6. **Do teachers face a heavy workload?**

   The amount of work is one of the main concerns in education. The good news is that since 2017, by law, the hours for preparing classes have increased from 25% to 35%. This means that if you are contracted for 44 hours per week, 28 hours will be for teaching and 16 hours will be for preparing and evaluating learning, which balances the workload.

7. **None, I am clear on everything.**
8. **No response**: See the non-response protocol.

Do you want information about any of the other questions?

- **Yes**: Return to the list of questions
- **No**: Go to MODULE 4
- **No response**: See no response protocol.

MODULE 4: INFORMATION ABOUT THE APPLICATION PROCESS TO HIGHER EDUCATION

Now I will answer possible questions you may have about the process of applying for higher education. Which of the following topics would you like information about?

1. **Dates for taking the University Entrance Exam and applying to universities**

   This year, the University Entrance Exam will be held between December 6th and 10th, by groups. Check your test date group on the DEMRE website (www.demre.cl). Results of the exam will be available on January 11th. From January 11th to 14th, you can apply to universities. This is done through the DEMRE and Mineduc application portal.

2. **Scholarships, financial aid, and application dates**

   Starting from October 5th, you can apply for student financial benefits to fund your career at www.fuas.cl. We recommend that you review the *Beca Vocación de Profesor* Scholarship, which finances 100% of your education degree if you score an average of 600 points between Verbal and Mathematics, or 580 if you graduate from a public or subsidized institution and are in the top 10% of grades in your high school. If you need more information about available scholarships, go to https://portal.beneficiosestudiantiles.cl.

3. **Requirements for applying to a teacher education program**

   To apply for an education program, you only need to meet ONE of the following requirements:

   - 500 average points between Verbal and Mathematics.
   - Average grades within the top 30% of your high school.
   - Having passed an access program to continue education studies in higher education recognized by Mineduc and having taken the PDT.

4. **How to choose a teacher education program**

If you need information about education programs, I recommend using the career search engine www.mifuturo.cl. There, you will find admission requirements, tuition fees, and curricula. To decide which education program to study, I recommend the following:

- Make sure to choose a program accredited for at least 4 years.
- Prioritize curricula that balance theoretical and practical courses, with practical experiences from the first year.
- Check that the graduate profile aligns with your interests and values.

5. **I have no questions**

   Thank you very much for your time. I hope I have resolved your doubts, and I will be available if you need my guidance another day. Best of luck, congratulations on considering becoming a teacher, and good luck with the University Entrance Exam!

Do you want information about any other topic?

- **Yes**: Go back to the list of topics
- **No**: Thank you very much for your time. I hope I have answered your questions, and I will be available if you need my guidance another day. Have a great day, congratulations on considering becoming a teacher, and good luck on the University Entrance Exam!
- **No response**: [Wait 1 day] Thank you very much for your time. I hope I have answered your questions, and I remain available in case you need my guidance another day. Have a good day, congratulations on considering becoming a teacher, and good luck on the University Entrance Exam!