Bureaucratic Training and State Capacity: Experimental Evidence from Peru's Judiciary [*]

Daniel Chen[†]   Manuel Ramos-Maqueda[‡]   Bernardo Silveira [§]

February 26, 2025

Abstract

We employ a randomized controlled trial to evaluate an intervention aimed at improving the quality of in-service training of public servants in the Peruvian judiciary. The intervention involves pedagogical specialists who remotely monitor, assess, and provide feedback to instructors of courses taught to active judges and prosecutors in the country's judicial academy. We find that the intervention significantly improves educational outcomes, including grades and students' satisfaction with the course. Furthermore, the intervention improves the quality of judicial proceedings of treated judges, increasing the ratio of verdicts per case and the rate of judges' attendance to litigant requests. This study demonstrates that a simple, scalable virtual intervention can effectively improve the training of bureaucrats and, consequently, enhance the quality of public service delivery.

[†]Toulouse School of Economics; email: daniel.chen@iast.fr
[‡]University of Oxford; The World Bank; email: mramosmaqueda@worldbank.org
[§]University of California, Los Angeles; email: silveira@econ.ucla.edu

1 Introduction

Professional training for public officials is ubiquitous. Every year, countries around the world provide in-service training to improve the skills and knowledge of civil servants—including police officers, judges, procurement and tax officials, doctors, nurses, teachers, and social workers—who deliver critical state functions. Such in-service training represents a significant public expenditure. For instance, in 2017 alone, the U.S. spent around $10 billion on training civil servants, accounting for about 4% of the annual budget for federal and state governments on personnel compensation and benefits (Credential Engine, 2021).

Despite the ubiquity and cost of programs to train public servants, there has been limited research on how to improve the quality and efficacy of this training.[1] Whereas recent studies analyze the addition of new topics to the education of public officials (Azulai et al., 2020; Banerjee et al., 2021; Mehmood et al., 2024), no empirical research to date has evaluated how to improve the delivery of existing classes. Such limited research stands in stark contrast with the vast literature evaluating the delivery of classes to school-age children, which provides a trove of evidence on the effectiveness of policies aimed at enhancing the learning of students (Glewwe and Muralidharan, 2016; Kremer et al., 2013; Evans and Popova, 2016). A natural—and overlooked— question then arises of whether at least some of the insights from the general education literature are transferable to the training of public officials.

In this paper, we help address this question by partnering with the Peruvian judiciary to experimentally evaluate an intervention at the Judicial Academy of Peru (AMAG), the country's premier institution for the instruction of civil servants in the justice sector. To our knowledge, this is the first randomized controlled trial to evaluate a judicial training intervention anywhere in the world. The intervention consisted of remotely monitoring, assessing, and providing feedback to instructors of an eight-month training program for active judges and prosecutors seeking promotion. Pedagogical interventions of this type are simple, scalable and have a documented track record of enhancing education in schools (Allen et al., 2011). We find that the intervention significantly improved the educational outcomes of class participants,

---

[1] The comprehensive review of empirical research on the personnel economics of the state by Finan et al. (2017) categorizes existing research into three key areas: selection, incentives, and on-the-job monitoring of public officials. This review, however, does not include any studies focused on public officials' training. Similarly, the systematic survey of field experiments in public administration by Hansen and Tummers (2020) includes no papers on the training of civil servants apart from daycare workers and school managers.

leading to higher grades and satisfaction with the course. More importantly, it raised the quality of judicial service delivery by treated judges in the year after the program, increasing both the ratio of cases reaching a verdict and the ratio at which judges attend litigants' requests.

Evaluating the training of public servants in the judicial setting in Peru is particularly important. The investment in a well-functioning legal system is a key element of state capacity, given its crucial role in contract enforcement and the promotion of private sector development (Besley and Persson, 2011). Accordingly, extensive evidence shows the critical role of the judiciary for economic growth and protecting citizen rights (Amirapu, 2021; Ramos-Maqueda and Chen, 2021; Kondylis and Stein, 2023; Chakraborty, 2016; Chemin, 2012; Laeven and Woodruff, 2007; Acemoglu and Johnson, 2005). However, as in many other developing countries, the judiciary in Peru is often perceived as inaccessible, unreliable, and inefficient. In a 2018 survey of Peruvian citizens, 68% of the respondents believed that most or all judges are corrupt, while 63% felt the same about prosecutors. Only 13% of the respondents reported having some trust in judges and magistrates, with the corresponding number for prosecutors being 16% (World Justice Project, 2022).

Our intervention included 604 active judges and prosecutors in an online training program from May to December 2020. The training consisted of nine consecutive courses of four weeks each. These courses covered various disciplines, from fundamental topics (e.g., ethics, judicial and legal interpretation), to specialized law subjects (e.g., civil, criminal, and administrative law), and complementary topics (e.g., managing judicial offices and case-based theory). Students were divided into 23 classes based on their professional position and geographical location. Within each course, half of the classes were randomly assigned to receive the treatment, while the rest were assigned to the control group. The intervention followed closely the Pre-Analysis Plan registered in the AEA RCT Registry.[2]

At the beginning of each course, AMAG informed instructors in treated classes about the intervention, consisting of the monitoring, assessment and provision of feedback to instructors. Specifically, a pedagogical specialist would attend parts of the first session of online lectures and assess the instructor's class delivery. Following the first session, the specialist would meet privately with the instructor to review the

---

[2]The pre-analysis plan (PAP) can be found online at https://doi.org/10.1257/rct.7113-1.0. This PAP encompasses two distinct interventions, although they are documented under the same registration. The current paper exclusively examines the first intervention, referred to in the PAP as a "teacher monitoring" intervention.

delivery of the class and offer feedback, covering topics such as student participation, case-based teaching, class organization and management, and the effective use of technology. The specialist would then join the second session of online lectures, scheduled for two weeks after the first one, and assess the instructor again. Instructors in control group classes did not receive any monitoring, assessment, or mid-course feedback. All classes in our study—regardless of treatment assignment—followed the same structure and format as any other class in the Judicial Academy.

The intervention improved the educational outcomes of students in treated classes in comparison to the control group. Specifically, we observe an increase of 0.12 standard deviations in the final grades of treated students relative to the control group, reflecting the enhanced academic performance of students. Importantly, tests and exams were the same for students taking courses in the same subjects, regardless of treatment status—which facilitates the comparability between the control and treatment groups. It is also worth mentioning that the grades in the program are meaningful to the students, as they constitute an explicit component in judiciary promotion decisions; Section 2.1 provides more details. Both judges and prosecutors benefited from the intervention, with a significant effect on homework, exam and final grades for prosecutors, and on reading grades for judges. In addition, students increased their satisfaction with the instructor and the class as a result of the treatment.

We then assess whether the treatment led to improved judicial service delivery outside the classroom. With this goal, we collect data on judges' professional performance in 2021, the year following the intervention, and match this data with judges who participated in the 2020 training program.[3] Due to the lack of publicly available data for prosecutors, this part of our analysis focuses solely on judges. Based on the available data, we build indicators of the productivity and the quality of judicial processes.[4] First, we find that treated judges are relatively more likely to reach verdicts in their cases rather than to end them for procedural reasons at earlier stages—which we interpret as evidence that our intervention fostered more in-depth consideration and investigation of the legal proceedings' matter. This improvement in the quality of judicial processes comes at no expense of productivity, as the treatment has no

---

[3]The data comes from the publicly available website "Conoce a tu juez", which includes productivity metrics of judges: https://sap.pj.gob.pe/casillero-digital/#/conoce-tu-juez.

[4]As explained in greater detail in section 3, we rely on publicly available data to measure judicial performance, which allows us to build a limited set of performance indicators. While the indicators we can evaluate with this data do not exactly match those in the PAP, they still allow us to measure the impact of the intervention on both productivity and quality metrics, which is the original intention described in the PAP.

impact on the judges' clearance rate (the ratio of cases resolved over cases filed).

To complement these results, we evaluate the impact of the intervention on the judges' engagement in a recently implemented program called "The Judge Listens to You," through which litigants can request appointments with their judges. This program aims to bring judges closer to citizens, improving their sense of proximity and access to the justice system. We find that treated judges substantially increase the likelihood of attending meetings requested by the litigants, reducing the no-show rates. After the intervention, the attendance rate of treated judges was 96% relative to the control group average of 68%. These results point towards a more attentive and better service for citizens as a result of the treatment.

Given the importance of these results, we further investigate the underlying mechanisms behind the intervention's effectiveness. The intervention consists of two main components—monitoring and feedback. Monitoring is the process whereby the specialist observes and evaluates the instructor's performance during both the first and the second class sessions. As for the feedback component, it involves the instructor receiving feedback after the first session to improve the delivery in the second session. While the monitoring component affects instruction in both sessions, feedback specifically targets improvements in the latter session. Although there is no clean causal design to separately identify the impact of the monitoring and feedback components, we would expect to see similar effects on class outcomes in both sessions if monitoring were the key mechanism for improved teaching. Conversely, if feedback were the main driver, we would expect improvements in the second session compared to the first.

To explore this question, we leverage the only two sources of data that are consistently collected during the first and second sessions: student satisfaction and instructor performance. Our findings indicate that while the treatment did not significantly affect student satisfaction in the initial session, there was a substantial increase in satisfaction in the subsequent session. Similarly, we find an improvement in the grades of treated instructors—as recorded by AMAG specialists—in the second session relative to the first. This improvement is present in all three dimensions of instructor performance: teaching skills, content mastery, and class structure. We find the largest effect on class structure, which includes the instructor's ability to effectively open and motivate the class, employ diverse teaching strategies, engage students, and provide an effective closing. When we correlate the change in student satisfaction with the change in instructor grading between the two sessions, we also find a positive and significant correlation, particularly for class structure. Collectively, these results suggest

that the feedback component drives the positive outcomes of the intervention, particularly by improving the structure of the class. Thus, our results point to the potential of feedback-oriented strategies to enhance the academic and professional outcomes of civil servants, opening the door for future research to explore this question in greater depth.

Our paper contributes to a burgeoning literature on the in-service training of civil servants. As evidenced by its omission in previous literature reviews, this type of training has, until recently, received limited attention from empirical research on the personnel economics of the state (Finan et al., 2017; Hansen and Tummers, 2020). Such a gap in the literature is especially striking in light of the evidence of large dispersion in the effectiveness of individual bureaucrats within government organizations, which points to the need for interventions to improve the lowest performing public employees (Best et al., 2023). Three recent papers have contributed to bringing attention to the importance of training in the context of civil service, showcasing how newly designed training modules can be effective at improving civil servants' performance. Banerjee et al. (2021) show that a three-day soft-skills training enhanced the performance of police officers in Rajasthan, India. Meanwhile, Azulai et al. (2020) evaluate the impact of a new productivity training module delivered through Ghana's Civil Service, and Mehmood et al. (2024) evaluate an empathy training delivered to junior ministers in Pakistan. Importantly, all of these studies assess the impact of newly created in-service training modules. In contrast, our research focuses on improving the delivery of *existing* training offered by the Judicial Academy of Peru, without providing new content. Our intervention is informed by the vast literature on the education of school-age children, which shows that policies aimed at changing pedagogical practices—e.g., through monitoring, the provision of feedback and basic advice to teachers, and the use of structured lesson plans—figure among the most effective means of improving student performance (Kremer et al., 2013; Araujo et al., 2016; Glewwe and Muralidharan, 2016; Evans and Popova, 2016; Piper et al., 2018; Angrist et al., 2024). Conducted in a remote learning environment, our intervention is both scalable and cost-effective, showcasing its potential for bolstering state capacity in settings where governance challenges and resource constraints persist.

Our study is also, to the best of our knowledge, the first randomized controlled trial on judicial training, as well as the first empirical paper to study judicial training in a developing country context. This is particularly relevant given the slow adoption of empirical methods in the legal field. The few existing studies on this

subject have focused on the United States, evaluating the integration of new topics into judges' curriculum (Ash et al., 2022; Baye and Wright, 2011). We instead focus on improving a training program provided by an under-resourced educational institution in Peru, a developing country marked by low confidence in the judiciary by the general public. Studying this question in a developing country is particularly policy-relevant. Over the years, international organizations have invested substantial amounts in capacity-building programs aimed at improving the training of judicial actors and strengthening the rule of law across developing countries. For example, USAID has historically allocated 25% of its justice reform assistance to judicial training (Hammergren, 1998), while over 50% of The World Bank's lending operations on justice reform have supported training programs (The World Bank, 2012). Despite such efforts and investments, there is very limited evidence on effective training interventions to improve judicial performance. Ours is one of the few empirical studies to date showcasing the potential of simple, cost-effective interventions to improve the performance of justice systems throughout the world (Sadka et al., 2024; Chemin et al., 2022).

Finally, our intervention is embedded in the concept of "Community of Practice" (CoP), which underscores the importance of shared learning, social interaction, and collective knowledge development in professional settings. Since the seminal article by Wenger (1998) introduced the concept, CoPs have been influential in the organizational and pedagogical literature and practice. A CoP is defined as a group of people who "share an interest for something they do and learn how to do it better as they interact regularly" (Wenger, 2011). Despite being widely used in bureaucracies (Cuddy, 2002), CoP has seen limited randomized control trial evidence. We provide, to the best of our knowledge, the first experimental evidence evaluating the impact of CoP-inspired pedagogical interventions in the civil service context. Our randomized controlled trial in Peru, which involved systematic instructor feedback and observation, demonstrated improvements in both educational outcomes (e.g., grades, satisfaction) and professional metrics (e.g., verdict rates, litigant engagement). These findings underscore the role of structured peer learning and feedback in enhancing judicial quality—a key determinant of institutional performance and economic outcomes. Our study uniquely integrates CoP principles into the economics of governance and bureaucratic performance, offering a scalable framework for improving public service delivery.

The remainder of the paper proceeds as follows. Section 2 discusses the back-

ground context and the study design. Section 3 presents the data sources and outcomes of interest. We then present our empirical results in Section 4. Section 5 concludes.

2 Context and Study Design

2.1 Context: The Judicial Academy Training

AMAG is a government agency in Peru dedicated to providing training to active and aspiring judges, prosecutors, and clerks. Each year, the Academy offers education to over 10,000 legal professionals. In this study, we partner with AMAG's Promotion Program (APP), which trains active judges and prosecutors seeking career advancement. The APP is a critical component of the promotion process for Peruvian judges and prosecutors. Completion of the APP is required for candidates to qualify for advancement, and the grade obtained in the program influences the promotion outcomes. Specifically, promotion decisions are partly based on a numerical score that depends on the applicant's APP grade.[5]

The APP offers three tiers of training tailored to the varying experience levels of legal professionals. These tiers correspond to promotion courses for levels 2, 3, and 4, aligned with the hierarchical positions of judges and prosecutors. Level 2 applicants include judges in lower courts and adjunct or assistant provincial prosecutors who aim to become specialized judges or provincial prosecutors. Level 3 applicants are specialized judges or provincial prosecutors aspiring to positions in high-level regional courts or public prosecution offices. Finally, level 4 applicants are high-level regional judges seeking promotion to the Supreme Court. As there was only one class of level 4 applicants in the period covered by our intervention, we exclude it from this study.

The 22nd APP was conducted from May to December 2020, offering nine courses (rounds) of four weeks for each level. The nine courses each cover a separate topic, categorized into three groups: fundamental topics (courses 1-4), specialized topics (courses 5-8), and a complementary course. Fundamental courses addressed ethics in the judiciary, legal interpretation and argumentation, constitutional law and appeals. Specialized topics split students into areas of law based on their specialty—generally civil or criminal law, with specific subspecialties such as family and administrative

---

[5]As per the Regulation for the Promotion of Judges and Prosecutors in Peru, the promotion process will consider the applicant's grade in the APP, "if the applicant has obtained a grade equal to or greater than 13 in the program that corresponds to the level to which she aspires" (Reglamento de Concursos para el Ascenso de Jueces y Fiscales, Junta Nacional de Justicia.)

law. These courses cover methodology for resolving cases, problems faced in the judicial proceedings, jurisprudence in these fields, or emblematic cases, among others. The complementary course dealt with judicial office management and included a two-week elective workshop. In addition, each of these courses took place in the midst of a transition towards case-based teaching in the Academy, with increased emphasis on practical applications of the lessons into the judge and prosecutors' day-to-day jobs. More detailed information on the topics of the program can be found in Appendix B.

Each of these courses lasts approximately four weeks. In the initial days of the course, participants are granted online access to class materials. Following the review of these materials, students have to complete online activities and assignments. They also attend day-long sessions on the second and fourth Saturday of each course, spanning from 9 a.m. to 5 p.m. At the end of each of these day-long sessions, students are required to fill out the satisfaction survey provided by the Judicial Academy. Throughout the course, students complete two tests and a final examination. More detailed information on a typical course schedule can be found in Appendix C.

The instructors of each course are university professors with a specialization in a legal discipline and a minimum of four years of postgraduate experience. The instructors are hired by the Judicial Academy to teach a specific course related to their specialization. The specific instructor thus changes every course, with very few instructors teaching more than one course per program.

2.2 Intervention: Instructor Monitoring and Feedback Program

From May to December 2020, AMAG rolled out an innovative intervention across the entire APP program. The intervention was randomized at the class level. Specifically, for each four-week long course, students were divided into 22 classes—with four classes for level 2 students and 18 classes for students in level 3. In the first course of the program, the 22 classes were randomly assigned into treatment and control groups, with stratification based on class level and location.[6] Thus, 11 classes—each with a different instructor—were treated, while the remaining 11 were in the control group. Students remained in the same treatment group during the majority of the courses, with some exceptions when they selected specialized courses in rounds 6-8. Out of the 604 students in total, 456 (representing 75.5% of the total) were fully treated or never treated, while the remaining students participated in some treated classes but

---

[6]While the instruction was online, students were grouped into different classes based on their location. This is why we also use class location as strata.

not in all of them (see Figure A1 in the Appendix).

Throughout the program, 97 classes received the treatment and 103 classes were in the control group.[7] Treated classes had instructors who were observed, assessed, and received feedback by a specialist from the Academy, while control classes did not receive any observation or feedback.

Each treated class had a unique instructor who was assigned a specialist. The specialist was a pedagogical expert already employed by the Judicial Academy, responsible for implementing the intervention in up to four classes per course. There were a total of six different specialists in charge of implementing the intervention throughout the entire program. These specialists also had prior teaching experience and involvement in guiding and monitoring instructional performance. They had been trained in constructive communication skills, enabling the monitored instructors to receive feedback positively and to facilitate improvements in their teaching processes.

The intervention consisted of two main components: instructor observation and assessment, and subsequent feedback to enhance performance. The pedagogical team at AMAG designed class evaluation guidelines, which included an "Observation Form" for standardized assessment of instructors, and a "Agreements and Commitments Form" for standardized feedback sessions. The latter form ensured that the discussion between the specialist and the instructor resulted in a set of agreements and commitments that the instructor would have to consider in their next class. The intervention was designed by methodological specialists at AMAG, based on evidence from the Education literature and on a review of similar experiences in other settings.

The timeline of the intervention for each course is as follows: The specialist first assessed the delivery of the class in the first class session, following the criteria in the "Observation form." After the first session, the specialist provided feedback to the instructor following the "Agreements and Commitments Form." Finally, two weeks after the first class session, the specialist assessed the delivery of the second session, following the same "Observation form" as in the first session. This timeline of events is summarized in Figure 2, which highlights in red the key components of the intervention.

During the first session, the specialist visited each class four times throughout the day, with each visit lasting an average of 15 minutes. Thus, the specialist spent a total of one hour in each class every day. During the visits, the specialist took notes,

---

[7]For ease of understanding, Figure 1 summarizes the structure of the program and the treatment assignment across classes and rounds.

Judicial Academy Training (JAT)								
May (Round 1)						June (Round 2)	...	Dec (Round 9)
Level	Total Students	Class #	Number of Students	Trainer Name	Monitoring Intervention			
2	157	1	32*	Trainer 1**	Yes (Mon A)***	...	...	...
		2	32	Trainer 2	No	...	...	...
		3	32	Trainer 3	Yes (Mon A)	...	...	...
		4	32	Trainer 4	No	...	...	...
3	457	5	25	Trainer 5	Yes (Mon B)	...	...	...
		6	25	Trainer 6	No	...	...	...
		...	...	...	...	...	...	...
		21	25	Trainer 21	Yes (Mon C)	...	...	...
		22	25	Trainer 22	No	...	...	...

*The numbers are based on a typical round. Some students did not take all classes, thus the total number of students is larger than the typical/average number of students in a class.

**Trainers change every round, each teaching 1 class per round up to 4 rounds.

*** There were 3 monitors in most rounds. 3 additional monitors assisted in some rounds.

Figure 1. Program structure

recorded in the observation form, which assessed the instructor in three main areas: (i) *teaching skills* (e.g., enthusiasm, student relationships, voice modulation, and language clarity); (ii) *content mastery* (e.g., the clarity of the presentation and the use of real life examples); and (iii) *class structure* (e.g., the use of motivational strategies and the evaluation of students). Appendix D provides details of the assessment indicators in the observation form.

In the week following the first session of classes, the specialist and instructor had a 30-minute feedback meeting to discuss and offer feedback on the teaching experience during the first session. During this meeting, the specialist first asked the instructor to perform a self-assessment, including their assessment of their own strengths and areas for improvement in the teaching skills, mastery of the content, and class structure. The specialist then shared their own feedback in terms of both the strengths and areas for improvement in those same topics. Based on this discussion, the instructor and specialist reached a set of agreements and commitments on how to improve the delivery of the class in the upcoming session of classes. All of this was documented in the "Agreements and Commitments Form," which was signed by both the specialist and the instructor. A sample of this form can be found in Appendix E.
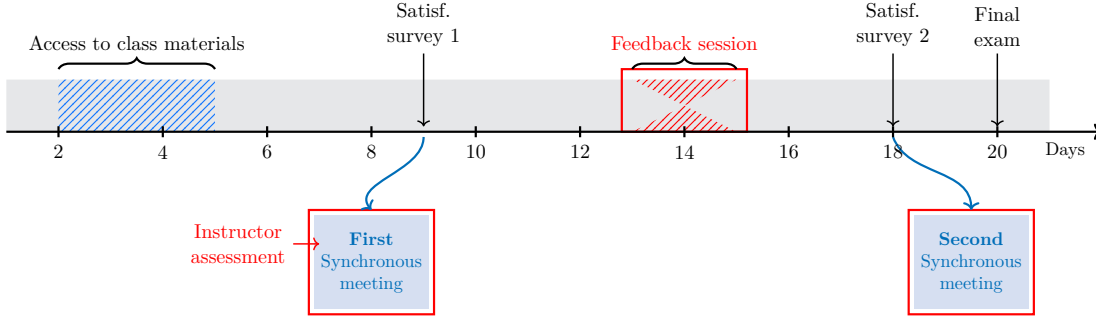
Figure 2. PCA - Calendar course

Table 1. Summary statistics and balance table at the class level

Statistics	Summary Statistics						Balance Tests	
	Treatment			Control				
Class-level stats	Mean	Sd	N	Mean	Sd	N	β	p-value
Number of students	26.06	4.05	97	25.72	3.83	103	0.01	0.40
Share of female teachers	0.22	0.41	97	0.19	0.40	103	0.02	0.53
Share of judges	0.32	0.18	97	0.31	0.19	103	-0.23	0.45
Share of prosecutors	0.68	0.18	97	0.69	0.19	103	0.23	0.45
Share of female students	0.33	0.11	97	0.41	0.09	103	-1.99	0.23
Age	46.34	3.23	97	45.45	3.24	103	0.03	0.80
Years of tenure	5.30	0.78	97	5.40	0.61	103	-0.14	0.52
Years in the bar association	17.86	3.00	97	17.62	2.62	103	-0.02	0.87
Share in criminal court	0.17	0.11	97	0.17	0.12	103	0.17	0.54
Academy's specialist female	0.66	0.48	97					

Note: This table presents balance tests on the monitoring treatment. We present summary statistics displaying means and standard deviations for treatment classes ("Treatment") and control classes ("Control"). Balance tests present an OLS regression on treatment, with strata (participant level and location) and round fixed effects.

3 Data

3.1 Academic performance and student satisfaction

There were 604 judges and prosecutors who enrolled in the APP. Most students took level 3 (72.77%), followed by level 2 (22.93%), while only 4.30% took level 4. As previously explained, we exclude level 4 from this study, since it had only one class per course. The proportions of prosecutors in levels 2 and 3 are 70.83% and 64.99%, respectively, with the rest of the participants being judges.

Administrative data on grades and student satisfaction was collected over nine rounds for all judges and prosecutors enrolled in the program. The grade data includes detailed information on homework, tests, and final exam grades for each round.

These tests and exams were the same for all students in the same course, regardless of the class treatment status. The satisfaction data consists of a series of Likert-scale questions about the student's satisfaction with their learning experience, the instructor's preparation, and the use of digital resources, among others. Students answered satisfaction-related questions twice within a round, right after each day-long class, and their participation in this survey is required by the Academy. These satisfaction questions are collected by the Academy before the final exam and before students receive their final grades in the course.

3.2 Professional performance

To evaluate the impact of the intervention on the professional performance of course participants, we employ publicly available data from the Judiciary of Peru's "Know Your Judge" (*Conoce a tu Juez*, in Spanish) platform.[8] This platform makes available information on several performance measures for individual Peruvian judges. We focus on judge-level measures referring to the professional performance in 2021—the year following our intervention—of judges who participated in levels 2 and 3 of the APP program in 2020. Since there are no data available on the professional performance of prosecutors, we restrict our analysis of professional outcomes to judges.

Performance measures available in the platform include the number of *verdicts* and the number of overall *decisions* per judge. Verdicts, in this context, constitute final decisions made by judges, based on the merit of the case. Meanwhile, the number of overall decisions consists of the number of cases decided for any reason by the judge; in addition to verdicts, it includes cases closed due to dismissal orders, final decrees, and other procedural reasons.

Also available through "Know Your Judge" is information on the judges' participation in a program titled "The Judge Listens to You" (*El Juez te Escucha*, in Spanish; henceforth, we refer to it as JLY). The program launched in October 2019 with the goal of bringing justice closer to the general public by fostering direct communication between litigants, their lawyers, and judges. Initially piloted in the civil and constitutional high courts of Lima, it was expanded to additional civil courts in the country in 2021. The program facilitates online meetings where participants can discuss procedural aspects of ongoing cases. Such discussions might address issues

---

[8]The following website (in Spanish) provides further information on the platform: https://www.gob.pe/14476-consultar-informacion-sobre-magistrado-conoce-a-tu-juez. We extracted the data used in our analysis in September 2023.

like sentencing delays, file assessments, or procedural flaws, without encroaching on substantive matters that are reserved for formal hearings to ensure all parties' rights are preserved.[9] The "Know Your Judge" platform contains, for each judge, information on the number of JLY meetings requested by litigants, as well as the number of meetings attended, canceled, or closed due to a no-show by the parties or the judge. To emphasize the importance of adhering to the JLY program, the platform explicitly lists the attendance of meeting requests as a performance metric of judges.

In addition to the data from the "Know Your Judge" platform, the Judiciary of Peru makes available case-level data with detailed information on court orders and other case-related events.[10] Using these data, we compute the number of cases filed in each judge's docket in the year 2021.[11] The case-level data does not include information on family cases. We, therefore, exclude judges specialized in family cases from the analysis relying on these data.

With the available data on professional performance, we construct indicators to assess both the productivity of judges and the quality of case proceedings.[12] We define below the exact outcome variables that we use in our analysis, organized across these categories.

Productivity: We measure productivity by computing the judge's *case clearance rate*, which is the ratio of cases resolved (*number of overall decisions*, as reported in the "Know Your Judge" platform) to cases filed (obtained from the case-level data). An increase in the case clearance rate is associated with higher judicial productivity.

Quality of proceedings: As a first metric for quality, we compute the *rate of verdicts*, defined as the ratio between the *number of verdicts* and the *number of overall decisions* (both statistics are from the "Know Your Judge" platform). We interpret a higher value as indicating a more in-depth investigation of the matter of the case,

---

[9]Per Administrative Resolution 000077-2021-CE-PJ, published in El Peruano.

[10]For further information on these data (in Spanish), see https://www.gob.pe/14192-consultar-casilleros-digitales-de-jueces-y-juezas. We extracted the data used in our analysis in September 2023.

[11]Unfortunately, the case-level data do not contain clear information on case resolution. We are thus unable to compute measures such as case length; the direction of judges' decisions; or even the number of decisions and verdicts by a judge in a given year, which we would have been able to compare with the metrics from the "Know Your Judge" platform.

[12]It is important to highlight that certain metrics initially proposed in the PAP, such as the case length, the fraction of decisions appealed or reversed, the length and direction of judicial decisions, or the textual measures of implicit bias, were not measurable due to data limitations (see footnote 11). Nevertheless, we were able to include in our analysis other valuable indicators—notably those derived from the JLY program, which was only implemented in large scale after our initial PAP submission in early 2020. Indicators related to this program provide relevant insight into the quality of judicial proceedings.

which leads to a final verdict. A lower value means that more cases ended due to procedural rather than substantive reasons.

We also measure the quality of proceedings by considering the judges' participation in the JLY program. Specifically, we define, for each judge, the *attendance rate* as the ratio between the number of JLY audiences successfully completed and the number of audiences requested by the litigants. A higher value of this rate indicates more engagement of the judge with the public, which we interpret as better service provision by the judge. Additionally, we consider, as a separate metric, the number of *meetings requests*—that is, the number of JLY meeting requests from litigants that the judge received. A higher value of this metric indicates greater demand for judicial interaction. To be sure, the interpretation of this metric is not as straightforward as that of the attendance rate. A large number of meeting requests can suggest that the litigants see the judge as more accessible, which is aligned with the mission of the JLY program. Conversely, it can also reflect an increase in procedural issues that require further discussion with the judge. In interpreting results related to the number of meeting requests below, we keep in mind this potential ambiguity.

4 Results

This section first presents the regression results for educational outcomes (grades and satisfaction) of all students—that is, both judges and prosecutors. Then, we evaluate the impact of our intervention on professional outcomes, with a focus on judges; as explained in Section 3.2, data on prosecutors' professional performance are not available. Finally, we assess the mechanisms behind our baseline results by focusing on two distinct components of the intervention—monitoring and feedback.

4.1 Grades and Satisfaction

To estimate the impact of the intervention on grades, we run the following regression specification:

$$Grades_{icr} = \alpha + \beta \times Treat_{cr} + \lambda \times \mathbf{Topic}_{cr} + \eta_r + \gamma \times \mathbf{Z}_{icr} + \varepsilon_{icr}, \qquad (1)$$

where $Grades_{icr}$ is a grade score for student i in class c during round r; $Treat_{cr}$ is an indicator that class c received the intervention during round r; $\mathbf{Topic}_{cr}$ is a vector

of dummies indicating the content of the course taught to class c in round r; η_r is a round fixed effect; and Z_{index} is a vector of control variables used in the stratification (namely, class location and participant level). The coefficient of interest, β, indicates the average impact of the intervention on grades. Standard errors are adjusted for clustering at the class level.

We proceed in a similar fashion to estimate the effects on satisfaction. Since there are two satisfaction surveys per round, we include the index s to account for the session. The regression specification is:

$$Satisfaction_{iscr} = \alpha + \beta \times Treat_{cr} + \lambda \times \mathbf{Topic}_{cr} + \eta_r + \theta \times Second_s + \gamma \times \mathbf{Z}_{imcr} + \varepsilon_{iscr}, \quad (2)$$

where $Satisfaction_{iscr}$ is a satisfaction measurement for student i in synchronous session s in class c and round r; and the remaining variables are as defined as in the specification for grades—with the addition of $Second_s$, which is a dummy indicating the second synchronous session in the round. As before, standard errors are adjusted for clustering at the class level.

Table 2 shows the main results on grades and satisfaction outcomes. We present the results for five evaluations taken by the students throughout the course (columns (1)-(5)) and the average satisfaction with the instructor and the course in the two class sessions (columns (6)-(7)). To facilitate the interpretation of the effects, we report estimates after standardizing each outcome based on the control group participants' mean and standard deviation.

First, we find an increase in final grades of 0.13 standard deviations relative to the control group mean. This final grade is a weighted sum of all other course grades, and, as a general assessment of the performance in the course, is the most important grade for students. Other grades, such as participation, reading, homework or exam grades, also change in the positive direction, but the effect is not statistically significant.

We then turn to evaluate the impact of the intervention on student satisfaction in columns (6) and (7). Satisfaction is reported after every full day of classes—and, therefore, there are two measures of it for every course. The Academy requires the students to fill in the satisfaction surveys, so we do not have attrition for these responses. We find significant increases in satisfaction with the instructor and the course—both in the order of 0.10 standard deviations.

A potential concern about the results in Table 2 is that the treatment status of the students in any round of our analysis is influenced by their class assignment in the first round of the program. To address this concern, Appendix Table A1 replicates

15

Table 2. Treatment effects on grades and satisfaction

	Grades					Satisfaction	
	(1) Forum	(2) Reading	(3) Homework	(4) Exam	(5) Final	(6) Teacher	(7) Course
Treatment	0.0686	0.0579	0.0891	0.1589	0.1263**	0.0993*	0.1004*
	(0.0766)	(0.0354)	(0.0544)	(0.0964)	(0.0608)	(0.0545)	(0.0524)
# Participants	604	604	604	604	604	604	604
Observations	5,033	5,053	5,090	5,067	5,094	10,099	10,099
R^2	0.13792	0.15050	0.11485	0.06360	0.08072	0.02618	0.02990
Round FEs	✓	✓	✓	✓	✓	✓	✓
Course FEs	✓	✓	✓	✓	✓	✓	✓

Notes: Standard errors in parentheses, clustered at the class level. All columns include strata controls (participant level and location), as well as round and course fixed effects. Columns 6-7 also include meeting fixed effects. The number of participants is the unique number of students, whereas unit of observation is participant-round in columns 1-5 and participant-round-meeting in columns 6-7. Minor variations in number of observations are due to some missing outcome data points for very few participants. Standard errors are clustered at the class level. Results are robust to alternative specifications. Grades and satisfaction outcomes are standardized with respect to the control group mean. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$

the analysis in Table 2, but with the standard errors clustered at the level of the students' class assignment in the first round. Reassuringly, we obtain results that are qualitatively identical and quantitatively very similar to those in Table 2.

Another potential concern about the main analysis in Table 2 is that, while instructors in the program teach at most one class per round, a relatively small number of instructors teach multiple rounds per year. It is possible that an instructor that receives the treatment in an early round of our intervention is later assigned to an untreated class. Thus, if the effect of the treatment on the instructor's performance lasts long enough, the exposure of the instructor to the treatment early on could affect the students' outcomes in subsequent rounds taught by the same instructor—regardless of their treatment status. To address this concern, we re-estimate our specifications from Table 2, restricting the sample to the class-round combinations that correspond to the first round taught by each instructor in the APP program in the year 2020. This subsample comprises roughly two thirds of all class-round combinations in our data. The results, reported in Appendix Table A2, are qualitatively the same as those obtained using the full sample. If anything, the estimated treatment effects are larger in magnitude; and the impact of the intervention on the reading, homework and exam

16

grades, which are not significant in Table 2, become significant once we focus on the subsample.

Appendix Table A3 presents the results for grades and satisfaction when the sample is split between judges and prosecutors. Despite the reduced sample size, which particularly affects the statistical power for judges, we still observe a significant increase in both grades and satisfaction for prosecutors, along with a notable improvement in reading grades for judges.

Taken together, our results on educational outcomes indicate that the intervention improved not only the objective performance of students in the class, but also their subjective assessment of the course. These findings are thus consistent with improved teaching by the instructors, learning by the student, and a better experience in the course overall.

4.2 Professional performance

As explained in Section 3.2, we measure the professional performance of judges in 2021, the year following our intervention. Since our experimental design randomized treatment assignment at the class level, and judges who participated in the APP program take multiple classes over the year, the degree to which judges were exposed to the treatmtment may vary from one judge to another. To leverage the variation in treatment intensity, we define the variable $Percentage\,Treated_i$ as the proportion of courses that received the intervention, out of the total of APP courses taken by judge i in 2020. We then consider the following specification to test whether the intervention had an impact on judicial performance:

$$y_i = \alpha + \beta_1 Percentage\,Treated_i + \gamma \times \mathbf{Z}_i + \varepsilon_i, \tag{3}$$

where y_i is the professional outcome indicator for judge i (which can refer to the ratio of verdicts, the judge's case clearance rate, the number of JLY meetings requested by litigants, and the attendance rate to those requests); and $\mathbf{Z}_i$ refers to the strata controls, which include location and participant level.

It is worth noticing that judges can either participate in panel decisions or in single-judge courts. Panel decisions are generally composed of 3 judges who decide on higher instance court cases. Meanwhile, single-judge courts are composed by a single judge who decides on the case, often at a lower instance court. In our results, we distinguish between two samples of judges: a comprehensive sample, which includes

17

judges who participated in both panel and single-judge cases throughout 2021; and a restrictive sample, which includes only judges who exclusively made single-judge decisions during the same period. This allows us to differentiate the potential effects of the intervention in cases where the judge is part of a team of judges relative to cases where the judge is the sole individual responsible for deciding in the case.

Table 3 reports the results from specification (3). In columns (1) and (2), the dependent variable is the judge's case clearance rate—with column (1) referring to the comprehensive sample of judges, and column (2) limiting the sample to judges who only decided cases in single courts. Regardless of the sample, the table shows positive but insignificant effects of the case clearance rate. Therefore, taken at face value, the results in columns (1) and (2) suggest that the intervention had no impact on the overall productivity of judges. But we believe that these findings should be interpreted cautiously, as the numerator and the denominator employed to construct the case clearance rate are taken from different data sources. Specifically, the number of decisions—the numerator in the clearance rate—is from the "Know Your Judge" platform; and we obtain the number of cases filed—the clearance rate's denominator—from the case-level data. As explained in Section 3.2, the latter data source does not include family cases. And, although we excluded from our analysis judges who specialize in family cases, the remaining judges might have family cases as part of their caseload—which would reflect in our measure of the number of decisions from "Know Your Judge." This disconnect between the numerator and the denominator in the clearance rate might lead to measurement error, which could affect the precision of the estimates reported in columns (1) and (2) of Table 3.

We now turn our attention to the effect of the intervention on the quality of judicial proceedings. The first such measure that we analyze is the rate of verdicts—that is, the number of verdicts relative to the total number of decisions taken by the judge. Columns (3) and (4) of Table 3 report the results for this dependent variable, with column (3) referring to the comprehensive sample, and column (4) restricting the sample to single-court judges. For both samples, we find that, in the year after the intervention, judges in the treated group increased their rate of verdicts very substantively; the increase is of 13.5 percentage points for the comprehensive sample, and 17.6 percentage points for single-court judges. Considering that the average rate of verdicts is roughly 32 percent for judges in the control group, these magnitudes imply increases in the rate of verdicts of treated judges by over 40 percent in the year following the intervention.

Table 3. Impact on judges' professional performance

	Rate of case clearings		Rate of verdicts		JLY attend rate	JLY aud. requests
	Dependent variable:					
	(1)	(2)	(3)	(4)	(5)	(6)
Percent. Treat	0.260	0.356	0.135**	0.176*	0.279**	50.171
	(0.299)	(0.543)	(0.056)	(0.096)	(0.136)	(34.819)
Dep Var Mean	0.49	0.47	0.32	0.33	0.68	53
Observations	149	76	158	80	44	44
R^2	0.020	0.022	0.081	0.116	0.295	0.140
Adjusted R^2	-0.029	-0.079	0.038	0.030	0.158	-0.027

Notes: All dependent variables are measures for the year 2021. Specifications (1)-(4) control for location and participant level; (5)-(6) control only for location. Table A4 shows results of specifications (5)-(6) controlling for location and participant level. The *rate of case clearings* is the number of cases closed relative to all cases filed. The *rate of verdicts* refers to the cases resolved with a verdict relative to the total of cases resolved by the judge. The *JLY attendance rate* consists of the proportion of requested audiences in the "Judge Listens to You" program successfully completed by the judge. The *JLY audience requests* refers to the total number of "Judge Listens to You" audience requests received by the judge.

As a complementary metric for the quality of proceedings, we consider the judges' participation in the JLY program, described in detail in Section 3.2.[13] In Table 3, column (5), the dependent variable consists of the judge's attendance rate to JLY audiences—that is, the number of completed audiences over the number of audiences that were requested by the litigants. The results indicate that the intervention increases the treated judges' attendance rate by 27.9 percentage points. Accounting for an average attendance rate of 68 percent among control-group judges, our findings imply that, in the year after the intervention, judges in the treatment group achieved an attendance rate close to 100 percent. Column (6) of Table 3 shows results in which the dependent variable is the number of JLY audience requests received by the judge. We find positive but non-significant effects of the intervention on these requests, indicating that the increase in the attendance rate discussed above is not due to a lower incidence of requests. That the point estimate is positive suggests that, if anything, treated judges obtain more requests than non-treated ones.[14]

---

[13]As explained in Section 3.2, the JLY program had not yet been extended to all courts in 2021; in particular, it had not been implemented in criminal and specialized courts. As a consequence, the sample of judges in our analysis of JLY participation data is smaller than the sample that we employ to evaluate the impact of our intervention on other professional outcomes.

[14]Moreover, in Appendix Table A4, we show that, if we introduce controls for participant level to

To summarize our results regarding professional outcomes, we find large, positive effects of the intervention on two measures of the quality of proceedings: the rate of verdicts, and the attendance rate to audiences in the JLY program. Both results suggest that treated judges behaved in a more attentive way towards the litigants. That is, treated judges were more willing to meet with litigants and hear their concerns and requests; and they were are also less prone to end cases for procedural reasons, thus giving to a larger proportion of litigants a chance to pursue their case until they reached a verdict. That these results are aligned strengthens the evidence that our intervention impacted the professional behavior of judges in the year following the experiment. Importantly, we find that the improvements in quality were not accompanied by a lower overall productivity of judges, as measured by the case clearance rate. Taken together, these results suggest that our classroom intervention translated into the practice of the treated judges, with a marked improvement in the quality of their decision making in court.

4.3 Mechanism: Personalized feedback

Given the substantial effects of the intervention on the academic and professional performance of students, in this section we ask: What component of the intervention drives the improved quality of teaching and student learning and satisfaction?

There are two main potential mechanisms of the intervention: the monitoring of the instructor by the specialist, and the feedback offered by the specialist to the instructor. The monitoring component consists of the observation and assessment of the instructor by the Academy's specialist. Instructors were informed at the beginning of the course that the specialist would monitor their teaching and provide feedback. Thus, we might expect that instructors—anticipating the monitoring—exert greater effort overall and improve their motivation and quality of teaching. The second main component of the intervention relates to the feedback meeting between the specialist and the instructor that takes place after the specialist observes the first session of the class. The feedback meeting allows both the instructor and the specialist to reflect on the strengths and areas of improvement in the teaching of the instructor. Furthermore, in these feedback meetings, the instructor also commits to improving

---

a specification like that in column (6), we obtain a positive and statistically significant coefficient for the effect of the intervention. As discussed in Section 3.2, it is not straightforward to interpret an increase in the number of audience requests, by itself, as a desirable outcome of the intervention. But the findings in the Appendix, at the very least, corroborate our conclusion that the increase in the attendance rate in Table 3, column (5) is not due to a smaller number of audience requests.

specific aspects of its teaching in the remainder of the course, as documented in the "Agreements and Commitments Form."

To examine the relative importance of the monitoring and feedback components in explaining our baseline effects, we leverage the fact that the instructor was observed in both class sessions, while the feedback meeting only occurred after the first session. The timing of the feedback meeting implies that it only affects the quality of teaching in the second session, whereas the monitoring component of the intervention presumably affects the quality of the teaching in both sessions. Therefore, if the feedback component were the main driver of the effects of our intervention, we would expect to see an improvement in the quality of teaching in the second class session, relative to the first one. Conversely, if monitoring were the chief component, the improvement in the quality of teaching would appear already in the first session.

The outcome variables needed to implement this exercise must be collected consistently during the first and second sessions of the class. Two types of variable in our data satisfy this requirement: those related to instructor performance, and those measuring student satisfaction. Instructor performance data were collected by the specialist, who observed and provided feedback to instructors in treated classes. As such, these data are only available for classes in the treatment group. In contrast, student satisfaction is available for all classes, allowing us to compare the evolution of the performance between the first and the second sessions for the treatment and the control groups.

Table 4 shows an increase in the grading of treated instructors' performance by the specialist in the second session relative to the first one for all areas of observation: teaching skills, content mastery, and class structure. The effect size is approximately double for the "class structure" grading component than for teaching skills and content mastery. As shown in the Observation Form in Appendix D, the class structure encompasses various elements, including how the instructor opens and motivates the session, the teaching strategies employed, student engagement and feedback, and the emphasis on key takeaways during the closing stage of the session. These elements see, on average, a greater improvement in the second session relative to teaching skills (e.g., teacher's enthusiasm, clarity, or voice modulation), or their mastery of the content and the topics discussed.

Table 5 presents the regression results for student satisfaction—both with the instructor (columns (1) to (3)) and with the course (columns (4) to (6)). The results indicate that there were no significant effects of the treatment on any of the

satisfaction measures in the first session (columns (1) and (4)), whereas positive and significant effects exist for both measures in the second session (columns (2) and (5)). Taking the difference between the satisfaction measures in the second and first sessions, we find positive and significant effects—with a magnitude of approximately 0.08 standard deviations both for satisfaction with the instructor (column (3)) and for satisfaction with the course (column (6)). These patterns are consistent with an improvement in the teaching quality in the second session, relative to the first one.

Finally, for classes in the treatment group, we also correlate the change in student satisfaction with the change in instructor grading between the two sessions in Table A5, in the Appendix. We would expect such a correlation to be positive if the increase in instructor performance were a driver of the change in student satisfaction between the first and second sessions. Indeed, we find that there is a significant association between the increase in instructor grading regarding class structure and the increase in student satisfaction with the teacher. This finding suggests that the improved class structure in the second session, relative to the first one, is associated with the observed increase in student satisfaction.

Collectively, the results in the present section suggest that the intervention was particularly effective at improving the quality of the teaching in the second session of the class. While our findings are consistent with the feedback component driving the effects of the intervention, we take them only as suggestive evidence, as we cannot rule out other potential mechanisms that could be behind the effect.[15] Additionally, we find that the improvement in class structure appears to drive the increase in student satisfaction and instructor grading. This result suggests that the intervention's success was due, at least in part, to its effect on more malleable aspects of instruction, such as class organization, motivational techniques, teaching strategies, student engagement, and the closing and review stages. In contrast, elements like content mastery or intrinsic teaching skills may be less influenced by our intervention. These results indicate the need for further investigation into the potential strategies that might be more effective for improving in-service training in settings like the one analyzed in our paper.

---

[15] For example, a competing hypothesis is that the monitoring component has a compounded effect, and becomes more effective with time—i.e., in the second session relative to the first one. While we consider this hypothesis less plausible, we cannot completely rule it out based on the available data.

Table 4. Instructor Grades per Session

Dependent Variable: Model:	Teaching Skills (1)	Content Mastery (2)	Class structure (3)
Variables			
session2	0.3871***	0.4724***	0.8608***
	(0.1026)	(0.1027)	(0.1300)
Fixed-effects			
Course	Yes	Yes	Yes
Round	Yes	Yes	Yes
Class	Yes	Yes	Yes
Fit statistics			
Observations	170	170	170
R^2	0.53233	0.46835	0.54794
Within R^2	0.05791	0.07118	0.21151

Notes: Each regression includes course, round, and class fixed effects. Clustered (Course) standard-errors in parentheses. Each of the columns represents a different outcome. Each outcome is computed as the average of all grading questions under each of these categories. The total number of points for each grading item is 20. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 5. Personalized feedback effects on satisfaction

	Satisfaction with teacher			Satisfaction with course		
	(1) First	(2) Second	(3) Diff.	(4) First	(5) Second	(6) Diff.
Treatment	0.0565	0.1423**	0.0837***	0.0589	0.1421**	0.0809***
	(0.0542)	(0.0567)	(0.0199)	(0.0520)	(0.0549)	(0.0194)
Number of Participants	604	604	604	604	604	604
Observations	5,052	5,047	5,045	5,052	5,047	5,045
R^2	0.02661	0.02904	0.00997	0.03096	0.03322	0.01254
Round fixed effects	✓	✓	✓	✓	✓	✓
Course fixed effects	✓	✓	✓	✓	✓	✓

Note: Standard errors are clustered at the class level. Columns 1 and 4 show coefficients for the first meeting subsample, Columns 2 and 5 show coefficients for the second meeting subsample, Columns 3 and 6 show coefficients for the difference between the first and second meeting. All columns include strata controls, i.e., location and participant level. Satisfaction outcomes are standardized with respect to the control group mean. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

5 Conclusion

The performance of judges and prosecutors is essential for a well-functioning justice system—a core component of state capacity. This is particularly important in developing countries such as Peru, where the justice system is seen as opaque and lacks citizens' trust. In this study, we partnered with the Judicial Academy of Peru (AMAG) to implement an RCT within the Academy's Promotion Program (APP), a program that trains active judges and prosecutors seeking promotion every year. We randomized whether instructors received an online monitoring and feedback program that aimed to improve the quality and effectiveness of their teaching.

We find that the intervention increased both class and job performance indicators. For judges and prosecutors in the treatment group, we observed an increase in satisfaction with the course and grades in the final exam. Furthermore, we find significant effects on the quality of judicial proceedings conducted by judges in the treatment group in the year after the intervention. In specific, we find an increase in the ratio of verdicts to any decision, which is also accompanied by an increase in the rate at which judges attended meetings requested by litigants. Our analysis suggests that the feedback provided to instructors played a significant role in driving these improvements in teaching quality and student learning.

To our knowledge, this is the first study to utilize a randomized experiment in the context of judicial training. The findings demonstrate the potential for enhancing the quality of education for judges and prosecutors, highlighting the positive impact that such improvements can have on the services provided to citizens. Given the substantial investments in judicial training by national and international organizations worldwide (USAID, World Bank), our results warrant further investigation in future research.

Additionally, our study contributes to the broader literature on state effectiveness and bureaucratic performance by showing how enhancing the quality of instruction for civil servants can improve the provision of public services. In contrast to prior research that evaluates newly-designed training modules (Azulai et al., 2020; Banerjee et al., 2021; Mehmood et al., 2024), we focus on enhancing the quality of the existing training of civil servants. Our results suggest that this approach, which has the potential to be applicable across various sectors beyond justice, leads to improved educational and professional outcomes for civil servants. Consequently, our research points to a promising new avenue for improving both the quality of instruction of public employees and public service delivery.

References

Acemoglu, Daron and Simon Johnson, "Unbundling institutions," *Journal of political Economy*, 2005, *113* (5), 949–995.

Allen, Joseph P, Robert C Pianta, Anne Gregory, Amori Yee Mikami, and Janetta Lun, "An interaction-based approach to enhancing secondary school instruction and student achievement," *Science*, 2011, *333* (6045), 1034–1037.

Amirapu, Amrit, "Justice delayed is growth denied: The effect of slow courts on relationship-specific industries in India," *Economic Development and Cultural Change*, 2021, *70* (1), 415–451.

Angrist, Noam, David K Evans, Deon Filmer, Rachel Glennerster, Halsey Rogers, and Shwetlena Sabarwal, "How to improve education outcomes most efficiently? A review of the evidence using a unified metric," *Journal of Development Economics*, 2024, *172*, 103382.

Araujo, M Caridad, Pedro Carneiro, Yyannú Cruz-Aguayo, and Norbert Schady, "Teacher quality and learning outcomes in kindergarten," *The Quarterly Journal of Economics*, 2016, *131* (3), 1415–1453.

Ash, Elliott, Daniel L Chen, and Suresh Naidu, "Ideas have consequences: The impact of law and economics on american justice," Technical Report, National Bureau of Economic Research 2022.

Azulai, Michel, Imran Rasul, Daniel Rogger, and Martin J Williams, "Can training improve organizational culture? Experimental evidence from Ghana's civil service," Technical Report, EDI Working Paper Series 2020.

Banerjee, Abhijit, Raghabendra Chattopadhyay, Esther Duflo, Daniel Keniston, and Nina Singh, "Improving police performance in Rajasthan, India: Experimental evidence on incentives, managerial autonomy, and training," *American Economic Journal: Economic Policy*, 2021, *13* (1), 36–66.

Baye, Michael R and Joshua D Wright, "Is antitrust too complicated for generalist judges? The impact of economic complexity and judicial training on appeals," *The Journal of Law and Economics*, 2011, *54* (1), 1–24.

Besley, Timothy and Torsten Persson, *Pillars of prosperity: The political economics of development clusters*, Princeton University Press, 2011.

Best, Michael Carlos, Jonas Hjort, and David Szakonyi, "Individuals and organizations as sources of state effectiveness," *American Economic Review*, 2023, *113* (8), 2121–2167.

Chakraborty, Pavel, "Judicial quality and regional firm performance: The case of Indian states," *Journal of Comparative Economics*, 2016, *44* (4), 902–918.

Chemin, Matthieu, "Does court speed shape economic activity? Evidence from a court reform in India," *The Journal of Law, Economics, & Organization*, 2012, *28* (3), 460–485.

_ , **Daniel L Chen, Vincenzo Di Maro, Paul Kieti Kimalu, Momanyi Mokaya, and Manuel Ramos-Maqueda**, "Data Science for Justice: The Short-Term Effects of a Randomized Judicial Reform in Kenya," Technical Report, Toulouse School of Economics Working Paper 2022.

Credential Engine, "Education and Training Expenditures in the U.S.," Technical Report, Washington, DC: Credential Engine 2021.

Cuddy, Colleen, "Cultivating communities of practice: A guide to managing knowledge," *The Bottom Line*, 2002, *15* (2).

Evans, David K and Anna Popova, "What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews," *The World Bank Research Observer*, 2016, *31* (2), 242–270.

Finan, Frederico, Benjamin A Olken, and Rohini Pande, "The personnel economics of the developing state," in "Handbook of economic field experiments," Vol. 2, Elsevier, 2017, pp. 467–514.

Glewwe, Paul and Karthik Muralidharan, "Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications," in "Handbook of the Economics of Education," Vol. 5, Elsevier, 2016, pp. 653–743.

Hammergren, Linn, "Judicial Training and Justice Reform," Technical Report, U.S. Agency for International Development 1998.

Hansen, Jesper Asring and Lars Tummers, "A systematic review of field experiments in public administration," *Public Administration Review*, 2020, *80* (6), 921–931.

Kondylis, Florence and Mattea Stein, "The speed of justice," *Review of Economics and Statistics*, 2023, *105* (3), 596–613.

Kremer, Michael, Conner Brannen, and Rachel Glennerster, "The challenge of education and learning in the developing world," *Science*, 2013, *340* (6130), 297–300.

Laeven, Luc and Christopher Woodruff, "The quality of the legal system, firm ownership, and firm size," *The Review of Economics and Statistics*, 2007, *89* (4), 601–614.

Mehmood, Sultan, Shaheen Naseer, and Daniel L Chen, "Altruism in governance: Insights from randomized training for Pakistan's junior ministers," *Journal of Development Economics*, 2024, p. 103317.

Piper, Benjamin, Joseph Destefano, Esther M Kinyanjui, and Salome Ong'ele, "Scaling up successfully: Lessons from Kenya's Tusome national literacy program," *Journal of Educational Change*, 2018, *19*, 293–321.

Ramos-Maqueda, Manuel and Daniel Chen, "The Role of Justice in Development: The Data Revolution," Technical Report, The World Bank 2021.

Sadka, Joyce, Enrique Seira, and Christopher Woodruff, "Information and Bargaining through Agents: Experimental Evidence from Mexico's Labour Courts," *The Review of Economic Studies*, 2024, *91* (46), 3677–3711.

The World Bank, "New Directions in Justice Reform: A Companion Piece to the Updated Strategy and Implementation Plan on Strengthening Governance, Tackling Corruption," Technical Report, The World Bank 2012.

Wenger, Etienne, "Communities of practice: Learning as a social system," *Systems thinker*, 1998, *9* (5), 2–3.

_ , "Communities of practice: A brief introduction," https://scholarsbank.uoregon.edu/items/6baf09ba-c319-4949-acef-cb220dc2953a 2011. Online; accessed December 12, 2024.

World Justice Project, "The Rule of Law in Peru: Key Findings from the General Population Poll 2022.," Technical Report, World Justice Project 2022.

A   Appendix: Supplementary figures and tables

- Figure A1: Distribution of participants by treatment intensity

- Table A1: Treatment effects on grades and satisfaction, with standard errors clustered at 1st round's class

- Table A3: Treatment effects on grades and satisfaction by profession

- Table A4: Treatment effects on meetings with litigants (Robustness)

- Table A5: Association between student satisfaction and instructor grades
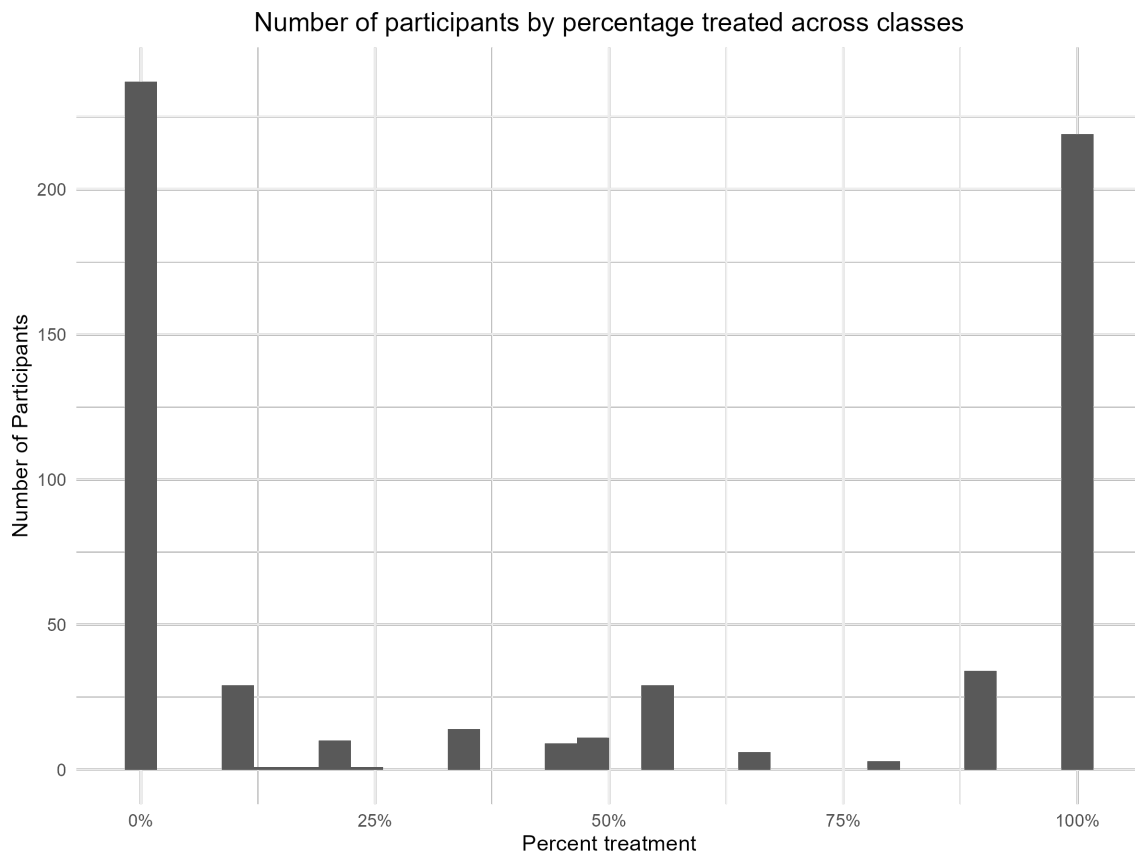
Figure A1. Participants by treatment intensity

Table A1. Treatment effects on grades and satisfaction, clustered at 1st round's class

	Grades					Satisfaction	
	(1) Forum grade	(2) Reading grade	(3) Homework grade	(4) Exam grade	(5) Final grade	(6) With teacher	(7) With course
Treatment	0.0686	0.0579	0.0891	0.1589	0.1263*	0.0993*	0.1004*
	(0.0777)	(0.0385)	(0.0542)	(0.0926)	(0.0648)	(0.0553)	(0.0574)
Number of Participants	604	604	604	604	604	604	604
Observations	5,033	5,053	5,090	5,067	5,094	10,099	10,099
R^2	0.13792	0.15050	0.11485	0.06360	0.08072	0.02618	0.02990
Round fixed effects	✓	✓	✓	✓	✓	✓	✓
Course fixed effects	✓	✓	✓	✓	✓	✓	✓

Notes: Standard errors in parentheses, clustered using the class they were assigned in the first round. All columns include strata controls (class level and location), as well as round and course fixed effects. Columns 6-7 also include meeting fixed effects.The number of participants is the unique number of students, whereas unit of observation is participant-round in columns 1-5 and participant-round-meeting in columns 6-7. Minor variations in number of observations are due to some missing outcome data points for very few participants. Standard errors are clustered at the class level. Results are robust to alternative specifications. Grades and satisfaction outcomes are standardized with respect to the control group mean. $* \ p < 0.10, \ ** \ p < 0.05, \ *** \ p < 0.01$

Table A2. Treatment effects on grades and satisfaction (Only first time teaching)

	Grades					Satisfaction	
	(1) Forum grade	(2) Reading grade	(3) Homework grade	(4) Exam grade	(5) Final grade	(6) With teacher	(7) With course
Treatment	0.1372 (0.1068)	0.0732* (0.0372)	0.1748** (0.0807)	0.2368* (0.1196)	0.2197** (0.0804)	0.1440*** (0.0442)	0.1344*** (0.0478)
Number of Participants	604	604	604	604	604	604	604
Observations	3,396	3,410	3,434	3,420	3,436	6,817	6,817
R^2	0.15727	0.17373	0.13182	0.08880	0.08943	0.02436	0.02764
Round fixed effects	✓	✓	✓	✓	✓	✓	✓
Course fixed effects	✓	✓	✓	✓	✓	✓	✓

Notes: Standard errors in parentheses, clustered at the class level. All regressions control for class lcoation and include round and course fixed effects.Columns 6-7 also include meeting fixed effects. These regressions only include the first time a teacher teaches a course.The number of participants is the unique number of students, whereas unit of observation is participant-round in columns 1-5 and participant-round-meeting in columns 6-7. Minor variations in number of observations are due to some missing outcome data points for very few participants. Standard errors are clustered at the class level. Results are robust to alternative specifications. Grades and satisfaction outcomes are standardized with respect to the control group mean. $* \ p < 0.10, ** \ p < 0.05, *** \ p < 0.01$

Table A3. Treatment effects on grades and satisfaction by profession

	Grades					Satisfaction	
	(1) Forum grade	(2) Reading grade	(3) Homework grade	(4) Exam grade	(5) Final grade	(6) With teacher	(7) With course
Panel A: Prosec							
Treatment	0.1013	0.0221	0.1286*	0.1909*	0.1535*	0.1723***	0.1568**
	(0.0807)	(0.0434)	(0.0740)	(0.1078)	(0.0874)	(0.0598)	(0.0566)
Participants	418	418	418	418	418	418	418
Observations	3452	3471	3495	3480	3499	6946	6946
R Squared	0.127	0.154	0.119	0.070	0.092	0.033	0.037
Panel B: Judge							
Treatment	-0.0134	0.1575***	-0.0133	0.0887	0.0607	-0.1048	-0.0598
	(0.0854)	(0.0444)	(0.0855)	(0.1007)	(0.0800)	(0.0762)	(0.0706)
Participants	186	186	186	186	186	186	186
Observations	1581	1582	1595	1587	1595	3153	3153
R Squared	0.176	0.161	0.146	0.073	0.088	0.041	0.034

Note: Standard errors are clustered at the class level. Grades and satisfaction outcomes are standardized with respect to the control group mean. Panel A shows regression coefficients for the prosecutor subsample. Panel B shows regression coefficients for the judge subsample. $*$ p < 0.10, $**$ p < 0.05, $***$ p < 0.01.

Table A4. Impact of Treatment on Judges' Attendance Rates

	Total Meetings	Attendance Rate
	(1)	(2)
Percentage Monitored	85.129*	0.305*
	(47.697)	(0.179)
participant_level_2	60.872	0.045
	(58.481)	(0.134)
Dep Var Mean	53	0.68
N	44	44
R^2	0.176	0.298
Adjusted R^2	-0.013	0.137

Notes: The regression controls for location and participant level strata. Total meetings refer to the number of meetings a judge scheduled during 2021. Attendance rate refers to the rate of meetings that the judge and litigant both attended during this same period. The number of judges in this regression is 44, as these are the sample of judges who participated in the pilot program of "El Juez te Escucha" out of the total sample of 193 judges.∗ $p < 0.10$, ∗∗ $p < 0.05$, ∗∗∗ $p < 0.01$

Table A5. Association between Student Satisfaction and Instructor Grades

Dependent Variable:	Content Mastery		Class Structure		Teaching Skills	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Constant	0.4102***	0.3933***	0.7181***	0.7021***	0.3792***	0.3589***
	(0.1415)	(0.1362)	(0.1403)	(0.1341)	(0.1223)	(0.1180)
Average Satisfaction	1.493		3.429*		0.1892	
	(1.996)		(1.978)		(1.725)	
Satisfaction with the teacher		2.100		4.219**		0.7487
		(1.972)		(1.942)		(1.709)
Fit statistics						
Observations	85	85	85	85	85	85
R^2	0.00670	0.01348	0.03493	0.05379	0.00014	0.00231
Adjusted R^2	-0.00527	0.00159	0.02330	0.04239	-0.01190	-0.00971

Notes: This table shows the results of a linear regression of the change in satisfaction between the second and the first session on the change in instructor grading between the second and the first session. "Average satisfaction" refers to the overall average of all responses to the satisfaction question, whereas "Satisfaction with the teacher" refers to the average of questions related only to satisfaction with the teacher. The different dependent variables refer to different dimensions that the instructor received grades on. The regressions shown do not include any controls. Results remain similar if fixed effects on class or round are included. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$

B   Appendix: Course Topics

Line of Training	Subject 2nd level	Subject 3rd level
Fundamental	**Course 1:** Fundamentals of Ethics in Judiciary	**Course 1:** legal argumentation
	Course 2: Constitutional Theory, Rights Fundamentals and Gender Approach in the administration of Justice	**Course 2:** Ethics in the Magistracy
	Course 3: Basic Fundamentals of Control conventionality and control constitutionality	**Course 3:** Control of Conventionality and Control of Constitutionality, Binding Precedents of the Constitutional Court and Standards of the IACHR
	Course 4: Interpretation and argument legal	**Course 4:** Emblematic cases of Law Constitutional Procedure
d	**Course 5:** • Methodology for solving criminal cases based on the theory of criminal law and the legal consequences of the crime • Theory of judicial decision in civil matters.	**Course 5:** Constitutional Framework of Law Administrative
	Course 6: • Problems in the Civil Process • Debatable Institutions in the Criminal Process • Problems in the Administrative Litigation Process	**Course 6:** • Assessment of the evidence at the appeal venue in the Civil Process • Assessment of the evidence on appeal in the Criminal Process

specialized	• Problems in the Administrative Litigation Process	Assessment of the evidence on appeal in the Criminal Process
	Course 7: • Analysis of jurisprudence and plenary agreements of the Supreme Court of Justice in civil matters • Analysis of jurisprudence and plenary agreements of the Supreme Court of justice in criminal matters.	**Course 7:** • Civil challenge law • Criminal challenge law
	Course 8: • Emblematic cases of Family Law • forensic investigation • Emblematic cases of Administrative Law	**Course 8:** • Plenary agreements of the Supreme Court of Justice in Civil Matters • Plenary agreements of the Supreme Court of Justice in Criminal Matters
complementary	**Course 9:** • Management and leadership of the Tax Office • Management and leadership of the Judicial Office	**Course 9:** Public Management: Skills managerial
	Workshop: Oral litigation and direction of hearings Workshop: Methods and techniques of investigation and case theory Workshop: Interculturality in justice peruvian	Workshop: Problems in Oral Litigation and direction of hearings Workshop: Role of the prosecutor and the judge in the era digital

C   Appendix: Course Schedule

COURSE SCHEDULE

INDUCTION WEEK

			Miércoles 14 de Oct. 20 20	Jueves 15 de otc. 2020	Viernes 16 de oct. 2020	Sbad 17 Oct 202
			START OF THE COURSE Revision of the guide didactics and syllabus of the course	Revision of the guide didactics, readings and case	Revision of the guide didactics, lectures and case	Review of material from study for the reading controls
Domingo 18 oct. 2020	**Lunes 19 oct. 2020**	**Martes 20 oct. 2020**	**Miércoles 21 oct. 2020**	**Jueves 22 oct. 2020**	**Viernes 23 oct. 2020**	**Sábado 24 Oct. 2020**
Review of material from study for the controls of reading	Video of Class N° 1 explanatory on the content of the course and component evaluative (45 minutes) Consultation chat N° 1 (45 minutes)	Review of material from study for the reading controls	Reading Control N° 1 (Start 00:00 hours)	Reading Control N° 1 (End 23:55 hours)	Video of Class N° 2 (Unit I and IL) (45 minutes) Consultation chat N° 2 (45 minutes)	1st. Synchronous session Unit I and II From 9:00 a.m. to 12:00 p.m. From 14:00 to 17:00

Figure A5. Program structure 1

Domingo 25 oct. 2020	**Lunes 26 oct. 2020**	**Martes 27 oct. 2020**	**Miércoles 28 oct. 2020**	**Jueves 29 oct. 2020**	**Viernes 30 oct. 2020**	**Sábado 31 Oct 2020**
Review of material from study for the forum	The discussion forum (Start 00:00 hours)	The discussion forum (End 23:55 hours)	Reading Control N° 2 (start 00:00 hours)	Reading Control N° 2 (End 23:55 hours)	Video of Class N°3 (Unit III) (45 minutes) Consultation chat N° 3 (45 minutes)	2nd Synchronous session Unit III From 9:00 a.m. to 12:00 p.m. From 14:00 to 17:00 Academic task
Domingo 1 nov. 2020	**Lunes 2 nov. 2020**	**Martes 3 nov. 2020**				
Review of material from study for the exam end	Final exam (Start 00:00 hours)	Final exam (End 23:55 hours) END OF THE COURSE				

Figure A5. Program structure 2

D Appendix: Observation form

The observation form includes three main sections: 1) Teaching skills; 2) Content mastery; and 3) Class structure. A summary of the characteristics in each section is the following:

1. Teaching skills: Assesses the teacher's enthusiasm, relationship with students, voice modulation, and clarity of language.

2. Content mastery: Evaluates the teacher's ability to clearly and comprehensibly present topics, use real-life examples and analogies, and emphasize key aspects.

3. Class structure:

   (a) "Opening Activities": assesses motivational strategies, communication of session objectives, and student engagement through questions, comments and/or activities.

   (b) "Intermediary Activities": examines the orderliness and relevance of exposition, the integration of case-based examples and related activities, the appropriate use of the syllabus and class materials, the orderly structure of class materials, and student evaluation and feedback.

   (c) "Closing Activities": focus on summarizing session content, evaluating learning outcomes, and providing feedback to students.

CLASS OBSERVATION SHEET
CLASSROOM 11
DATE 16.5.2020

1. **ACADEMIC PROGRAM** : 22 PCA
2. **Course or workshop** : LEGAL ARGUMENTS
3. **Teacher** : XXXX XXXX
4. **Learning Unit** :
5. **Date of synchronous/asynchronous sessions of the course or workshop** : 16/05/2020
6. **Observed session date** : 16/05/2020
7. **Start and end time when the observation is performed** : 9:00 - 13:00 AND 14:00-18:00
8. **Time (start, development, or closure) at which it was observed (indicate what was observed)** : It is indicated according to the lesson plan found in the Guide, indicate what will be observed, according to the Plan.
9. **Monitor or companion** : XXXX XXXX

INSTRUCTIONS
1. The duration of the observation will be 60 minutes (01 chronological hour).
2. Have the lesson plan, which is in the Didactic Guide.
3. The monitor (companion) must observe the actions of the teacher and assign an assessment according to the following scale:

SCALE	DESCRIPTION
A = 20	Outstanding
B = 18	Achieve the expected objectives with satisfaction
C = 15	It is in process.
D = 12	Requires support to achieve goals
E = 8	Presents serious difficulties or can do it and does not

4. If at the time of supervision any of the items cannot be observed, leave blank.

TEACHING SKILLS (HD)	A (20)	B (18)	C (15)	D (12)	And (08)
1. It transmits interest and enthusiasm towards the class topic.		x			
2. Relates to students in a cordial and respectful manner.	x				
3. Its volume, timbre, pitch and voice modulation are adequate.		x			

		A (20)	B (18)	C (15)	D (12)	And (08)
3.	Carry out the didactic actions according to the times programmed in the Class Plan (See didactic guide)	x				
Process of reflection on what has been worked on						
4.	It promotes the reflection of the students in the class on what has been worked on relating it to their jurisdictional work through questions or activities.	x				
Academic guidance (attention to queries)						
5.	Guide consultations in a clear and timely manner.	x				
Teaching strategies used.						
6.	Uses the case method or other teaching strategy (brainstorming, role play, directed debate, among others) for the development of the content or learning activities of the session.	x				
7.	Develop cases or other activities that allow to strengthen the subject worked.	x				
8.	Follow the procedure for the use of the teaching strategy (case method, brainstorming, role play, directed debate, among others)		x			
Study materials (syllable, teaching guide, or others) and resources						
9.	Use the syllable and the learning activity tutorial for the development of your class.	x				
10.	Use resources (videos, whiteboard, or other) to facilitate learning	x				
Organization of work in the classroom						
11.	Perform individual and/or group activities in the learning session.	x				
12.	Give clear indications to carry out the work whether individual or group.	x				
13.	Provides adequate and timely feedback to students, when performing some group or individual work, pointing out the difficulties encountered and positively valuing the successes.		x			
Learning assessment and feedback						
14.	Evaluates students through the use of cases or other activity that allows to see the application of what they learned in class to their jurisdictional work.	x				
15.	Use rubric or homework analysis sheet or other assessment instrument to grade the work of the students.	x				
16.	Provides adequate and timely feedback to the students, pointing out the difficulties encountered and positively valuing the successes.		x			
Average HC = (Total AD/16)=19.6						

CLOSURE ACTIVITIES (AC)	A (20)	B (18)	C (15)	D (12)	And (08)
1. Perform closing activities: ask the class group to indicate ideas strength of what was developed in the session, and / or evaluate the results through a final exam or the elaboration of an academic task, etc.		x			

		A (20)	B (18)	C (15)	D (12)	And (08)
4.	The language used is clear and intelligible.	x				
Average HC = (Total HD/4) = 19						

CONTENT MASTERY (DC)	A (20)	B (18)	C (15)	D (12)	And (08)
5. Develop topics in a clear, orderly and understandable manner.	x				
6. It uses examples from the work context and relates them to the contents developed.	x				
7. Explain the topics using examples, analogies, etc. that facilitate understanding.	x				
8. It emphasizes the most important aspects of the subject he works.	x				
Average HC = (Total DC/4)=20						

MOMENTS OF THE SESSION -FORMATIVE PROCESS (M)

INITIATION ACTIVITIES (AI)	A (20)	B (18)	C (15)	D (12)	And (08)
Motivation and initiation					
1. Perform a motivational activity at the beginning of class (anecdote, example, dynamics, video, etc.) related to the topic of the learning unit you will be working on.					
2. Communicates the capacity to be achieved in the session relating it to the usefulness for its jurisdictional work.					
3. It collects and uses the previous knowledge (knowledge they may have regarding what is addressed) of the students to build the new learning, through previously selected activities and strategies.					
4. Encourage the student to participate constantly through clear questions, comments, or activities.	x				
Average HC = (Total AI/4)=20					

MOMENTS OF THE SESSION -FORMATIVE PROCESS (M)

DEVELOPMENT ACTIVITIES (DA)	A (20)	B (18)	C (15)	D (12)	And (08)
Development of the learning session					
1. The presentation is orderly, precise and relevant.	x				
2. The selected examples and/or cases are addressed according to the level (I, II, III or IV) that corresponds to them, if applicable, and the place (Lima or province) where they work.	x				

		A	B	C	D	And
2.	It uses assessment instruments to rate the student's learning.					
3.	Provides adequate and timely feedback to the students, pointing out the difficulties encountered and positively valuing the successes.	x				
Average HC = (Total AC/3)=19						

CONCLUDING REMARKS
Very well the performance of the teacher in the development of the session, the initiation activities should be considered to optimize their performance.

E Appendix: Instructor feedback

The design of the intervention goes beyond simply observing instructors and ensuring they do their job. Personalized feedback, in addition to monitoring, facilitates the sharing of experiences and the targeted improvement in the delivery of the class. Below is an example of the type of feedback instructors receive and the actions instructors agree to take to improve their teaching.

Instructor Feedback Example Strengths:

- Highlighting the teaching skills transmitting interest and enthusiasm about the subject, the respectful and cordial relationship with the students, the volume, timbre, tone and modulation of the voice, and the use of clear and intelligible language.

- Mastery of the subject is highlighted: the use of examples from the labor sector to facilitate understanding and the linkages made with the content taught.

- The development of the initial activities: the presentation of the activities to be carried out.

- The presentation is clear, precise and relevant.

- Provides adequate and timely feedback to students.

Agreements / commitments to improve:

- The use of the camera when there is student participation to ensure human interaction and meaningful learning.

- Clarification in the reasoning. State the skill that will be addressed in the session.

- Carry out closing activities: ask the class for the main ideas from the session and the instructor complements the class contributions.

- Use some of the tools on the Blackboard platform.

ACADEMIA DE
LA MAGISTRATURA

ACT OF AGREEMENTS AND COMMITMENT

1. ACADEMIC PROGRAM		: TRAINING PROGRAM FOR PROMOTION
2.	course or workshop	: Plenary Agreements of the Supreme Court of Justice in Matter Civil
3.	teacher	▆▆▆▆▆▆▆▆▆▆
4.	Assigned classroom	: 30
5.	Observed learning unit: Learning Unit No. 01 and 02	
6.	Observed session date	: Saturday, October 24, 2020
7.	Date and time of beginning and end in which the feedback and the Minutes	: October 29, 2020 8:00 p.m. - 8:30 p.m.
8.	Monitor or companion (Specialist in the area of Methodological Design)	▆▆▆▆▆▆▆▆▆▆

1. SELF-ASSESSMENT OF THE TEACHER

Strengths:
- Motivation in the students for the participation and attention to the synchronous session.
- The teaching experience of the teacher in synchronous work means that the sessions can be be more dynamic.

Aspects to improve:
- Improve time management in student participation.
- Promote the participation of all students.

2. SCOPE OF WHAT WAS OBSERVED

Strengths:
· The teacher has a lot of management of information sources. · The teacher's
experience in the development of the topics covered means that the students can answer all their questions with the course teacher.

Aspects to improve:
- Increase motivation in the subject matter, to achieve more participation of the entire cluster.
- In order to improve access to information sources, the search for readings will be promoted. more updated.

ACADEMIA DE
LA MAGISTRATURA

3. AGREEMENTS AND COMMITMENTS

AGREEMENTS	COMMITMENTS
• The teacher agrees that the most active participation of all magistrates registered in the classroom should be promoted. • The teacher will promote the use of more sources of information.	· In the next session, we will seek to motivate the students more in order to achieve the participation of the majority of the students. · It will seek to take more care of the participation times of the students, so that everyone can participate.

OBSERVATIONS:
- · The teacher proposes that: During the development of the course, all the jurisdictional plenary sessions could be analyzed and, in particular, analysis of the last plenary session. Because of its importance.
- · Regarding the X Full Casatorium on the Examination, the teacher proposes that a special academic course or activity should be carried out in order to further promote their knowledge among magistrates and magistrates

_____
(Methodological monitoring team)

_____
activity teacher

F Appendix: Email to announce the treatment to the instructor

Email sent to each instructor in the Promotion Program of the Judicial Academy who taught classes in the treatment group:

Dear Dr. [Name],

Instructor of the Legal Interpretation and Argumentation Course,

I would like to take this opportunity to extend a cordial greeting, thank you for your valuable participation as an instructor in the Promotion Training Program, and inform you that, within the framework of the Regulation of the Professional Regime that provides teaching services to our institution and the Pedagogical Innovations being implemented by the Academy of the Judiciary, an impact evaluation of this training process is being carried out this year, with the support of the World Bank and our methodologists. This activity includes the monitoring and methodological support of the teaching work in the courses that are part of the 22nd APP, with the aim of providing guidelines and methodological orientations to help instructors achieve optimal performance in synchronous sessions, ensuring academic excellence and service quality.

In this regard, the methodological monitoring team of the Academic Directorate of AMAG will enter your videoconference room in the two synchronous sessions, and we request your valuable support in this matter. After this, they will contact you to provide feedback on your performance.

I would appreciate it if you could facilitate the development of both activities (monitoring and feedback), all with the aim of optimizing your performance as an AMAG instructor.

We wish you much success.

Sincerely,

Hipólito M. Rodríguez Casavilca

Academic Director, Judicial Academy of Peru