

Classroom Quality and Teacher characteristics in Ecuador

Yyannú Cruz-Aguayo
Nicolás Fuertes
Sara Schodt

Social Sector

TECHNICAL
NOTE N°
IDB-TN-01755

Classroom Quality and Teacher characteristics in Ecuador

Yyannú Cruz-Aguayo
Nicolás Fuertes
Sara Schodt

October 2019



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

Felipe Herrera Library
Cruz-Aguayo, Yyannú.

Classroom quality and teacher characteristics in Ecuador / Yyannú Cruz-Aguayo,
Nicolás Fuertes, Sara Schodt.

p. cm. — (IDB Technical Note ; 1755)

Includes bibliographic references.

1. Teacher effectiveness-Ecuador. 2. Teachers-Ecuador-Attitudes. 3. Teacher-student relationships-Ecuador. 4. Classroom management-Ecuador. I. Fuertes, Nicolás. II. Schodt, Sara. III. Inter-American Development Bank. Social Sector. IV. Series. IDB-TN-1755

<http://www.iadb.org>

Copyright © 2019 Inter-American Development Bank. This work is licensed under a Creative Commons IGO 3.0 Attribution-NonCommercial-NoDerivatives (CC-IGO BY-NC-ND 3.0 IGO) license (<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) and may be reproduced with attribution to the IDB and for any non-commercial purpose. No derivative work is allowed.

Any dispute related to the use of the works of the IDB that cannot be settled amicably shall be submitted to arbitration pursuant to the UNCITRAL rules. The use of the IDB's name for any purpose other than for attribution, and the use of IDB's logo shall be subject to a separate written license agreement between the IDB and the user and is not authorized as part of this CC-IGO license.

Note that link provided above includes additional terms and conditions of the license.

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.



Classroom Quality and Teacher characteristics in Ecuador

Yyannú Cruz-Aguayo, Nicolás Fuertes, Sara Schodt

Abstract

This project analyzes teachers' practices for approximately 2000 teachers in 204 schools in Ecuador. Each year a group of teachers was filmed and screened using the CLASS instrument to measure their practices. The results show that teachers perform better on Emotional Support and Classroom Organization domains and that the results are consistent across grades. Also, the project analyzes the impact of teachers' characteristics and skills on their CLASS score. Being a tenured teacher and having more experience are strongly related to CLASS score. Also, there is a strong relationship of IQ score, inhibitory control and attention and some traits of the Big Five, and their practices in the classroom.

Key words: Teacher quality, teacher characteristics

JEL codes: I21, I24

1. Introduction

For over six years now, the Inter-American Development Bank has run a longitudinal randomized controlled trial in Ecuador called Cerrando Brechas (Closing Gaps), using a multi-year random-assignment experiment to see how classrooms, teachers, peers and parents affect learning outcomes. The project has collected groundbreaking information on the relationship between teaching quality and children's learning. The project's primary motivation has been to better understand the characteristics or practices of the teachers who are most successful in closing the achievement gap between the poorest children in their classroom and their better-off schoolmates. One of the most powerful tools that the Bank has employed during this time is the Classroom Assessment Scoring System (CLASS) and this technical note will focus both on the instrument and its application in Ecuador in 5 different grades (Kindergarten-4th grade) between 2012 and 2016.

Targeted early childhood education interventions have proven to be an effective tool for improving school performance and life chances for poor or disadvantaged children in the United States; a growing body of evidence suggests that the same is true in Latin America (Burchinal et al., 1997; Campbell & Ramey, 1995; Lazar et al., 1982; Schweinhart et al., 2005; Peisner-Feinber et al 2001; Halpern 1986). Also, there is evidence that shows that the interactions with adults and peers have positive impacts on children's development and learning process (Fox et al., 2010)

Since children's academic and social abilities are largely dependent on the *quality* of their experiences in the classroom, it follows, then, that this aspect of the classroom should be assessed. For this reason, classroom quality is measure using the Classroom Assessment Scoring System (CLASS) which has been validated in several contexts and countries¹ and is one of the most promising instruments to measure classroom quality (Pianta et al., 2008). Besides measuring classroom quality, it is also useful to measure three different domains related to it: Emotional Support, Classroom Organization and Instructional Support (Pianta et al., 2008).

In terms of defining quality, previous attempts to measure classroom quality have traditionally concentrated mainly on infrastructure, physical inputs like materials, and curriculum and on improving teacher qualifications or credentials, or reducing class size (Hanushek, 2003). Recent meta-analysis of the effectiveness of hundreds of such interventions across programs and spanning

¹ Besides in Ecuador for this project (Araujo et al., 2016), the CLASS has been applied in Australia, Chile, China, Portugal (Tayler et al, 2013; Leyva et al, 2015; Hu et al, 2016; Cadima et al, 2010).

all grades, however, shows that there is very little evidence to suggest that these aspects of classroom experience have any significant impact on children's achievement or social competence (Hanushek, 2003; NICHD & Early Child Care Research Network, 2002). In fact, a review of the majority of such interventions suggests that increasing inputs or making teacher qualifications more stringent has little or no impact on children's outcomes in their early years of schooling (Hanushek, 2003; La Paro et al., 2004; La Paro et al., 2009).

So, what impacts children's academic and social outcomes? A number of different studies have found that children in high quality classrooms, as measured by the quality of their *interactions* with each other and with their teachers, not only initiate and participate in higher-order learning activities with their peers, but also achieve at higher levels on academic preparedness evaluations (Campbell & Ramey, 1994; Greenburg et al., 2001; Hamre & Pianta, 2005; Howes & Hamilton, 1993; Kontos & Wilcox-Herzog, 1997; Phillips et al, 2000). In one study, Mashburn et al. (2008) collected data on the quality of 671 pre-K classrooms, assessing quality by looking not only at infrastructural and curricular design inputs, but also at interactive processes. They also examined connections between these two types of quality (inputs and interactions) and child achievement outcomes (Mashburn et al, 2008). While the study found relatively little association between teacher qualifications, presence of physical materials or curricular design and child outcomes, it did find a strong association between classroom climate, as demonstrated by aspects of teacher-child relationships, and children's social and academic development (Mashburn et. al., 2008). In addition, high rates of positive interactions between kindergarten teachers and their students have been shown to be significantly associated with lower rates of "bad" behavior and higher socio-emotional competence in children as they progress through the first grade (Pianta & Nimetz, 1991).

Despite a growing body of evidence suggesting that teacher-child interactions are one of the most important components of quality and one of the most important predictors of children's cognitive growth, well-known measures of classroom quality such as ECERS, ECCRN, OCRE or ECCOM place less importance on teacher-child interactions and more heavily weight infrastructural or curricular organization aspects of the classroom (La Paro et al., 2004; La Paro et al., 2009).

The document is organized as follows. Section 2 describes the CLASS instrument, its dimensions, reliability to measure teacher quality, and some countries where it has been applied. Section 3 describes the application and coding process of the CLASS in Ecuador. The results and main

findings are described in Section 4 and, finally Section 5 presents the conclusions based on the results obtained.

2. The CLASS Instrument

The CLASS observation tool was developed by a team of researchers at the University of Virginia to evaluate the quality of teacher-student interactions in preschool and elementary classrooms (Pianta et al., 2008). Simply put, CLASS is based entirely on interactions between children and their teachers. For children in preschool through 3rd grade, the CLASS measure provides a common metric to describe how teachers utilize the materials they have available to them and how they interact with their students, considering three major areas (termed “domains”) of classroom characteristics: Emotional Support, Classroom Organization, and Instructional Support (Pianta et al., 2008). In the context of developing countries, the alternatives to CLASS mentioned above discriminate heavily against schools with few material resources, without fully accounting for the fact that empirical evidence backs the claim that real and high-quality classroom experiences are not dependent on the mere presence of such resources. The CLASS instrument, on the other hand, provides an opportunity to isolate the principal mechanism through which children learn and develop socially in the classroom context (Connor et al., 2004; Greenberg et al., 2001). Figure 1 gives an overview of the three major domains and the ten corresponding dimensions used by CLASS to define and assess classroom quality.

As an example, Emotional Support Domain is one of the three broad categories into which interactions between teachers and students can be grouped. As with the other two domains, Emotional Support, as it relates to children’s emotional and social expressions in the classroom, is backed by substantial literature identifying it as an important indicator of school success (Blair, 2002; Denham & Weissberg, 2004; Raver, 2004). The evidence shows that being in a classroom with appropriate emotionally supportive interactions is related to children’s socio-emotional development process. In particular, the literature has found that emotionally supportive interactions have positive impacts on peer behavior along the school year (Perry et al., 2007), peer aggression (Merrit et al., 2012) and academic outcomes (Rudasill et al., 2010).

Classroom Organization domain relates to classroom routines and teachers proactiveness. The evidence shows that being on a classroom that better manages behavior and attention has a positive impact on the time students spend on-task and they are also able to better regulate their attention (Rimm-Kaufman et al., 2009). There is also evidence on cognitive and academic

development, particularly on literacy gains (Downer et al., 2010; Ponitz et al., 2009). Finally, Instructional Support Domain is related to promoting order thinking and providing quality feedback. The literature has found a relationship between this domain and students' positive academic outcomes (Burchinal et al., 2008; Hamre & Pianta, 2005; Mashburn et al., 2008), particularly academic skills such as language, reading and math (Burchinal et al., 2010).

Scores for each dimension range from 1 (minimally observed) to 7 (frequently observed)² and represent the extent to which that dimension is characteristic of the classroom. Observation can begin at the same time the school day begins, or at any other predetermined time arranged with the teacher. Classroom observation is done in cycles, with each cycle consisting of a 20-minute period of uninterrupted classroom observation, where observers follow closely the “who, what and how of everything that happens at the classroom level, with particular attention to the teacher’s instructional interactions and behaviors,” while jotting notes on their CLASS coding sheets. The observation is followed by 10 minutes for CLASS enumerators to review their notes and record their codes on the same sheet.

The CLASS has also been approved for use with videotaped classroom observation. The concern with coding from video is of course that the camera may not capture the entire range of classroom interactions, so Pianta et al. (2008) suggest a number of guidelines for accurate and high-quality footage, including the following: teachers and students must be informed in advance that they will be filmed, and given the chance to ask questions and raise concerns; the video should run continuously from one activity to another, including transitions; and the camera should be mounted on a tripod and focused primarily on the teacher, but also pan the students to capture their expressions and experiences. In addition, Pianta et al. (2008) recommend filming from the side rather than the back of a classroom to better capture the expressions of both the teacher and the students.

It should be noted that as CLASS observers are evaluating a classroom or video, their scoring process for each dimension is not a merely running down a checklist of presence or absence of certain behaviors or indicators, but a holistic and composite description of classroom experience. This means that even classrooms receiving a high score of 7 on some dimension, like *Positive Climate*, may not display all of the indicators of high positive climate outlined in the CLASS Manual,

² For the dimension Negative Climate, the scoring system is reversed, with 1 indicating the absence of negativity and higher scores indicating higher levels of negativity (Pianta et al., 2008).

as long as other consistent indicators of positive teacher-pupil relationships are present (Pianta et al., 2008). After at least four observation cycles have been completed for a single classroom, composite scores for each dimension across the cycles are calculated by averaging the scores for each cycle. Similarly, averages of the scores from each dimension can be averaged to arrive at a composite score for each domain.

As a side note, it is important to note as a general comment on CLASS, that since the measure assigns a single score to an entire classroom of children, rather than an individual score reflecting each child's experience, the instrument assumes that each child in a given classroom experiences a similar level of quality so that their classroom quality CLASS score correlates with their individual achievement scores.

2.1. CLASS as a Professional Development Tool

The CLASS is also emerging as a powerful professional development tool. It has long been known that teachers are the most important in-school factor affecting a child's learning, but there are few tools outside of student test scores to evaluate teacher effectiveness; CLASS offers a way to address the issue by helping teachers identify and model the types of interactions known to improve children's emotional and cognitive development.

A large-scale professional development effort using CLASS was begun by the National Center on Quality Teaching and Learning (NCRECE). For this randomized control trial, teachers were assigned to a course aiming two goals: i) understand that every single interaction with children are foundational for their learning process; ii) provide information about the type of interactions that are good for children's language development. This course was based on the CLASS framework and the results show that teachers that attended the course used better interactions than the ones that were in the control group (Hamre et al., 2011).

One final note on using CLASS in professional development is that this designation of the instrument requires a distinct set of skills from those used for monitoring and evaluation purposes. While both activities require a close working knowledge of the CLASS instrument, CLASS observers are only trained to capture the quality of a classroom environment over a specific period in time, while CLASS professional development staff are experts in helping teachers to use the information from CLASS observations to reflect on their own practice and improve the quantity and quality of those interactions that are fundamental to optimizing student outcomes.

3. CLASS in Ecuador

The Inter-American Development Bank applied CLASS for the first time in 2011 in Ecuador for the longitudinal study *Cerrando Brechas* (Closing Gaps) a longitudinal randomized controlled trial collecting groundbreaking information on the relationship between teaching quality and children's learning from kindergarten through elementary school. The project's primary motivation was to better understand the characteristics or practices of the teachers who are most successful in closing the achievement gap between the most disadvantaged children in their classroom and their better-off schoolmates.

The company hired by the Inter-American Development Bank to carry out the CLASS work began by recruiting candidates for filming in Ecuadorian classrooms based mostly on their previous experience doing field work. The majority of the group had little academic training beyond high school and no prior technical knowledge of cameras or filming, but each had demonstrated a good work ethic in previous projects in terms of being reliable (showing up to work every day at the right time) and resilient (able to work long hours in fairly arduous conditions).

The filmers participated in a two-day classroom training with two instructors, one expert in the CLASS instrument and in the project, and the other a professional filmer and film professor. The participants spent the first day learning about the project, the CLASS instrument, and the general protocols that they would be expected to follow in the field; they also extensively role-played interacting with teachers, explaining their presence in the school, and dealing with resistant, nervous, or overly talkative teachers and/or children. The second day was entirely dedicated to technical training about the camera and its controls. Participants were trained on simple, tripod-mounted, handheld GZ-MG630 cameras, with a memory card for up to 30 continuous hours of filming, set to automatic control and to record at the lowest quality setting (resulting in smaller and more manageable electronic files). At the same time, a group of filming field supervisors was trained to download the video material to individual laptops and external hard drives, which would be done every day after filming, and to simultaneously review the quality of the material filmed and provide feedback to the filmers at the moment of download.

Filmers spent the next four days practicing filming in schools in and around Quito, Ecuador. The first day, most of the participants only filmed one hour of classroom time, during which each was closely supervised by the technical supervisor and given a lot of feedback and troubleshooting help. Later in the afternoon at the training headquarters, they all viewed each other's footage and

informally evaluated it in a whole group setting. During each of the following three days each participant filmed an entire school day and later watched their footage in a whole group setting back at the training headquarters in the afternoon, with input from their supervisors and fellow training participants. The last day, the filmers and supervisors remained at the training headquarters reviewing more footage, discussing challenging situations, and once again practicing the download of video material. The supervisors then reviewed the videos from the last day of filming and selected a final group of filmers for the project, based on the quality of their work over the training period and on their participation and dedication to the project as observed during the training period.

3.1. Filming in the Field

The classroom filmers worked every day and often filmed two school shifts (one in the morning and one in the afternoon) per day. Their arrival was not announced to the teachers whose classes they would be filming in with any prior notice, although they were accompanied at all times by a supervisor and provided with an official letter of entry from the Ministry of Education that they presented to the school principal before entering the school. Each filmer worked in a team of two with a supervisor who helped them each get set up in their respective classroom before the beginning of the school day and provided technical support throughout the day as needed. As an initial quality check, the supervisor also reviewed the videos each evening while they were downloaded and saved onto a laptop computer and an external harddrive.

At the beginning of each filming session, the filmer either said or wrote on a piece of paper the identifying information for the school and classroom, and all videos were saved as .mp4 files with the same information. Specifically, during the first round of filming, the filmers spoke this information into the camera, but in subsequent rounds of filming, the filmers wrote the date, name of the school, classroom information, teacher's name, and their own name on a piece of paper which they filmed for a few seconds at the start of the video of the school day.

3.2. Editing/Segmenting:

The video material was edited by two trained technicians in Ecuador who had no former experience in editing but had proven adept at repetitive-task jobs (like data entry) and at using the computer. The two technicians participated in a half-day training on the edition process. They follow a strict and consistent protocol with all videos, starting the segmentation process at the end

of the first hour of the video (the decision having been made to omit the first hour of all videos to allow for the novelty of the camera to wear off for the children and teachers) and making sure that all segments contained at least 5 children and their main teacher during at least 15 of the 20 minutes that the segment ran. The second segment starts where the first one ends, following the same procedure, and the third and fourth also follow the same protocol; in theory, they can begin as soon as the previous segment ends. However, if there is a period of more than 5 accumulated minutes during any 20-minute stretch in which either the minimum number of children or the main teacher are not present in the video, then editors are instructed to run the tape until they find the first available 20-minute stretch of video that meets those conditions. Additionally, none of the segments may contain more than 5 minutes of recess and/or lunch/snack time. All segments are labeled with the same identifying information as the unedited material, in addition to their respective segment numbers 1-4.

3.3. Coder Recruitment

The applicant pool for the CLASS observers was recruited through the listservs of a number of different academic departments (economics, education, psychology, sociology, and anthropology) at local universities and through some word-of mouth inquiries made by the company hired to execute the project. Candidates were screened individually for interest and ability to commit long-term to the project, and a final group of 14 candidates was then invited to participate in a week-long CLASS observer training.

3.4. Coder Training

The CLASS observer training and all subsequent work were supervised by a Teachstone-certified CLASS K-3 trainer. The trainer is also a CLASS pre-k and a CLASS Toddler Teachstone-certified observer, and prior to her work on this project had two years of extensive coding experience observing Chilean videos as part of her work with Harvard University's *Un Buen Comienzo* project in Santiago, Chile. The training in Ecuador followed the Teachstone CLASS observer training model but lasted slightly longer because the trainer required that the participants read the entire CLASS manual while in class, as an official part of the course training.

During the initial training, participants trained on official CLASS video clips to learn to recognize and correctly classify behaviors within each of the ten CLASS dimensions and were guided through the coding of a full-length CLASS video in English to see an example of a highly-scored classroom

environment. The remainder of their training and reliability testing took place with Ecuadorian videos. Prior to actually working with recruits, Spanish-language materials were prepared in a week-long workshop. This involved the development of sample videos, shot in Ecuadorian classrooms and master coded by two bilingual, CLASS certified trainers and a third bilingual CLASS observer, all of whom were familiar with Ecuadorian culture. During the master coding workshop, all of the videos were coded individually and then the group extensively discussed and documented evidence for each dimension, eventually arriving at a consensus about the CLASS score.

3.5. Coder Certification

At the end of the training week, candidates were given two days to study their manuals, notes, and the practice videos that they had watched in the training, and then they were given a reliability exam consisting of 5 master-coded Ecuadorian CLASS videos. A final group of 8 observers (coders) and 1 full-time back-up was selected; those selected 1) coded the certification test videos with 80% reliability using the CLASS standards and 2) showed aptitude, inquisitiveness, and motivation during the observer training week. The group then spent an additional two days coding videos together to gain additional confidence and practice. During these sessions, each coder watched and then coded the same video individually, and then discussed his or her codes as a group with the CLASS trainer.³

3.6. Coding Process

The selected group of Ecuadorian CLASS observers works 5 days a week and codes 9 video segments, randomly assigned and from different classrooms, each day. Each day begins with a group reliability session (described below), after which coders individually code 8 video segments. Coders work no more than 4 hours a day on those segments, are required to take a 15-minute break at least once during the coding and are encouraged to take breaks as frequently as they need to stretch, get water, etc. The coders are supervised during their entire shift by an administrative supervisor who sits in the coding room with them to ensure that they are not sharing information and that the coding sheets turned in at the end of every video are completely and accurately filled

³ In the subsequent year of coding we lost a few of the members of the original team and so a smaller group of only 6 candidates was selected to participate alongside already-certified CLASS observers in a CLASS Toddler training. All candidates participated in a similar training and certifying process to the K-3 experience, and from that group, a new team of two new observers were selected to complete the team of 4 previously-certified observers. Since that time, the team has retained its composition of 6 observers.

out. The administrative supervisor also deals with any and all administrative and technical issues that may arise.

The group reliability video with which each day begins is prepared by the CLASS supervisor. The group video chosen is usually one that has been flagged during the double-coding process. All videos are double-coded by two independent observers and those videos with large differences in their CLASS score are flagged for a third coding by a third independent observer. Based on a preliminary analysis of CLASS data from Ecuador that revealed a breakdown of high-and low-variability dimensions, videos are coded a third time if they have a difference of more than 1 point in the dimensions of Negative Climate, Concept Development, Quality of Feedback, or Language Modeling; if there is a difference of more than 2 points in any of the other dimensions they are also flagged for re-coding. In general, these videos are also the ones that have the most complex situations and so are the most useful for group discussions. When the coders arrive to work in the morning, they each individually watch and code the group reliability video, and meet with the CLASS supervisor to discuss each dimension in detail, documenting their observations at the indicator level and referring each score back to the descriptions in the Manual, to arrive at an agreed set of codes for the video. The CLASS supervisor facilitates the discussions, making sure that everyone is both participating in and understanding the process.

After the group reliability coding, the group then spends anywhere from a half hour to an hour and a half discussing their work from the previous day. As background to this discussion, each coder makes notes of complicated situations or videos during the course of their daily work and this information and the corresponding videos are then sent to the CLASS supervisor at night to review and prepare for discussion the following day. After watching the videos, the supervisor creates some discussion points regarding the interesting cases to share with the whole group. The following day, the group talks about these issues in the abstract and uses the Manual to discuss how such scenarios might be coded and which dimensions are most relevant to them. These discussions remain relevant to the observers in their ongoing work, because the curriculum changes throughout the course of the school year, presenting them with new classroom situations and different coding challenges.

When the group members begin their individual coding work, the CLASS supervisor stays connected to answer superficial questions that may arise during the process (for example, “A child spilled water on the ground, and a cleaning lady came into the classroom for a few minutes to

clean it up; do I register her as a second adult in the class?” or, “This video has 5 minutes and 20 seconds of recess time, do I still code it?” Or, “I saw a child steal something from the teacher’s desk, and the Manual doesn’t refer to that kind of situation. What do I do?”). The CLASS supervisor does not give scoring advice, but does direct coders to re-read certain sections of the Manual.

3.7. Coding Reliability

Reliability control happens in a number of ways: first, to keep track in a very general way of any observers who seem to be generating a lot of third codings, the supervisor receives a report of the videos which must be coded a third time every week, along with the information of which observers were responsible for the scoring. In cases where a pattern seems to be emerging, the supervisor pulls observers from the coding and works with them individually to re-train them, re-testing them before returning them to work. In Ecuador, this occurred only at the very start of the coding process, and when the two individuals who were pulled from the coding team failed to make sufficient improvement after a re-training and re-testing period, they were dismissed from the team altogether, leaving only 6 full-time observers. The second way that reliability happens is that every morning, each coder must hand in an individually coded scoring sheet before discussing the video from the group reliability coding session, providing the supervisor with additional information to identify coders who are consistently coding unreliably. Specifically, coders who code more than 2 videos unreliably during the course of a week would in theory be pulled from the team for re-training and re-testing, but this has not yet happened, because no observer has been flagged as coding consistently unreliably according to the aforementioned criteria.

All observers participate in a re-training course and take a new certification test at the start of each cycle of coding, which is generally anywhere from 6 to 10 months at a time.

Finally, in addition to the administrative supervisor, the CLASS supervisor, and the 6 full-time CLASS observers, the team has 1 back-up member who participates in the group reliability coding sessions every day, but only codes full-time if another team member is unable to work.

3.8. Verification

To ensure that all of the teachers filmed in 2011 baseline were the same teachers filmed in the later 2012 data collection, all of the schools were visited prior to filming to confirm the presence of the teacher and her classroom information. Additionally, once the filming took place in 2012, all of

the videos from 2011 were compared with the videos from 2012 to make sure that the labeling information was correct and that the teachers were indeed the same in both years.

4. Findings

4.1. CLASS results

Figure 2 shows univariate densities of the distribution of total CLASS and its three dimensions for teachers in the study by grade. Panel A shows the total CLASS distribution where most of the teachers have a score in the medium range (3-4) but there is still room for improvement since some teachers have lower scores. In Panel B the distribution of the Emotional Support of the CLASS shows that almost every single teacher is in the medium range (3-5) which indicates that teachers show some of the desired behaviors for this dimension. Panel C shows that the Classroom Organization dimension is the one where teachers are better score since some of them get higher scores (5-6). This means that this dimension is the one where teachers are showing more desired behaviors. Finally, Panel D shows the distribution for the Instructional Support dimension which is the most difficult domain and the one where teachers get lower scores. Every single teacher in the study is scored in the lower range (1-2) meaning this is the domain where teachers need more support. These results are also consistent with the ones find in Araujo et al. (2016).

Figure 3 shows the mean of the total score and each domain by grade. As in Figure 2, Figure 3 shows that the Instructional support is the one where more support is needed, whilst the Classroom Organization domain is where teachers show more desired behaviors and better practices. It is important to note that the above results are consistent across grades and among different groups of teachers⁴. Figure 3 shows that the mean across grades is very similar even though the teachers are different. This is an interesting result because the study filmed different groups of teachers in 5 different years⁵, but the results remain very similar. Since the results are very similar across grades, it allows to get some information about the practices and behaviors of teachers in this specific sample of schools (204 approximately), particularly on their teaching practices and the domains where teachers need more training in this schools.

⁴ Teachers changed every single year and no same teacher could participate in the study twice so in every grade/year a whole new group of teachers was filmed.

⁵ The grade filmed in each year was different. In 2012, Kindergarten teachers were filmed, in 2013 first grade teachers, in 2014 second grade teachers, in 2015 third grade teachers and in 2015 fourth grade teachers.

4.2. Teacher's characteristics and CLASS results

The most interesting aspect in the data collected is that it allows to estimate the impact that different teacher's characteristics have on their CLASS scores and its domains in different grades. For Kindergarten to 4th grade teachers (around 2000 teachers), information about their tenure, different measures of experience (time teaching at school, time working as a teacher and a dummy variable of whether or not the teacher has 4 or more years of experience), sex and age was collected. Furthermore, for the Kindergarten teachers (approximately 400) sample there is also information about 3 different tests applied to them: Big Five test (Personality traits), Stroop test that measures executive function and the Wechsler Adult Intelligence Scale (WAIS) IQ test.

Table 1 summarizes the characteristics of teachers in the sample for all grades. On average, teachers are in their mid-40s and they most of them are women (87%). Around three-quarters are tenured, and they have approximately 18 years of experience. In order to estimate the impact of these characteristics and others for the teachers in the panel dataset the following regression was estimated:

$$y_{it} = \beta_0 + \beta_1 age_{it} + \beta_2 sex_{it} + \beta_3 \bar{x}_{it} + \delta_t + e_{it}$$

where y_{it} represents teacher i score in CLASS (total and each one of its domains) in grade t , age_{it} , sex_{it} represent teacher i age and sex in grade t , \bar{x}_{it} represent different combinations of teacher i characteristics (tenure and measures of experience) in grade t , δ_t represents a fixed effect of grade t , and e_{it} represents the error term.

Table 2 shows the results of the model for the total score and each of the domains. First, column 1 suggest that tenured teachers are the ones that obtain a better CLASS score, particularly they obtain a score between 0.19 and 0.38 SD more than the teachers without tenure depending on the analyze score. Second, regarding the different measures of experience there are two important results: i) It is more important the overall experience as a teacher rather than the experience in the specific school. In particular, the effect of one more year of experience is between 0.012 and 0.025 SD on the score, while the effect of one more year of experience in the school is 0.009 and 0.019 SD (Columns 2 and 3); ii) The dummy of 4 or more year of experience turns out to be an important one meaning that rookies tend to have lower scores in the CLASS score. For example, in column 4 having 4 or more years of experience results on a 0.40 SD higher score in the total CLASS.

Finally, the results are robust to including every single variable (tenure dummy, measure of experience, age, sex and the year fixed effect) in the model as shown in columns 5-7.

These results show that, in general, there is a strong relationship between teachers' characteristics and their practices as measure by the CLASS instrument. Being a tenured teacher and having more experience seem to be the more important factors to understand the way teachers score in the CLASS and their practices in the classroom. Also, it is important to underline that these characteristics are more strongly related to the Emotional Support and the Classroom Organization domains.

It is also important to estimate the impact that different skills, measure with three different tests (Big Five, Stroop and WAIS IQ), have on the CLASS score for a cohort of Kindergarten teachers. For this reason, the following regression was estimated:

$$y_i = \beta_0 + \beta_1 age_i + \beta_2 sex_i + \beta_3 \bar{x}_i + e_i$$

where y_i represents teacher i score in CLASS (total and each one of its domains) in Kindergarten, age_i , sex_i represent teacher i age and sex in Kindergarten, \bar{x}_i represent different combinations of teacher i results in the tests (Big Five personality traits⁶, Stroop test⁷ and WAIS IQ⁸) in Kindergarten, and e_i represents the error term.

Table 3 shows the results for each score and its domains and their relationship with the tests applied to the Kindergarten teachers. Column 1 in Panel A shows that there are two personality traits of the teachers that impact their CLASS score. On the one hand, an increase in one SD in the Neuroticism score decreases the total score in 0.14 SD, while in the other hand an increase in one SD in Agreeableness increases the total CLASS in 0.09 SD. These results imply that teachers with less emotional stability underperform in the CLASS but teachers that are more friendly and cooperative perform better. Column 2 shows that there is also a relationship between the score in the inhibitory control and attention test and the total score, pointedly a 0.13 SD effect. This means that the teachers that obtain a better score, show more ability to control their attention and

⁶ The personality traits in which the test focuses are neuroticism, extraversion, openness, agreeableness, and conscientiousness. The test consists of 48 questions for each trait (total of 240 questions). Respondents should answer each question on a 5-point, Likert-like scale (where 1 is "totally disagree" and 5 is "totally agree").

⁷ The Stroop test is a test designed to measure attention and inhibitory control where teachers were shown a set of incongruent stimuli and they were supposed to inhibit the automated response. For example, they were shown the word "blue" printed in green ink and they were asked the color of the ink.

⁸ In the study the Spanish-speaking version of the Wechsler Adult Intelligence Scale (WAIS-III) was applied.

impulses. Column 3 shows that the teachers that get a higher score in the IQ test also obtain a higher score in the CLASS (0.19 SD effect). Finally, in column 4 the results suggest that both the Neuroticism and the IQ test effects are robust to including every single score in the model. Panel B of Table 1 shows the results for the Emotional Support Score and the results are very similar to those found for the total score.

Panel C presents the results for the Classroom Organization Score. Column 1 shows that this score is also negatively correlated with the Neuroticism, but the coefficient is slightly smaller than the one for the total score (0.09 SD vs 0.139 SD). More important on this specific measure of the CLASS, the Extraversion of the teachers has a positive effect (0.09 SD). This means that the more confident and assertive teachers are the ones that manage their classroom in a better way. Also as shown in Panels A and B there is a positive relationship between the CLASS domain and the score of the teacher in the inhibitory control and attention test, and the IQ test. Finally, Column 4 shows that most of the results are robust to including every single test score in the model, except the executive function one. Finally, in Panel D the results are very similar for the Instruction Support Score, except the Neuroticism one. However, it is important to note that in Column 1 there is a small impact of Openness and the domain score. This implies that the most creative teachers score better in this domain.

5. Conclusion

In general, the results show that teachers in Ecuador are doing some things well and the type of practices they are engaging in are effective, particularly the ones related to Emotional Support and Classroom Organization. However, teachers in Ecuador need support in order to get better practices related to the Instructional Support domain. Also, the results show that the results are similar no matter the grade they teach at and they are consistent for all the grades that were filmed during the project.

This technical note analyzes the impact of different teachers' characteristics and skills for teachers in 5 different grades in Ecuador that were filmed and screened using the CLASS instrument. For this purpose, the note uses two different samples of teachers (pool dataset using teachers from Kindergarten-4th grade and a sample of just Kindergarten teachers).

The results show that teacher characteristics are important to understand their school practices, particularly their experience and whether they are tenured are the most important factors to

understand their CLASS scores. Moreover, the results show that not only teachers' observable characteristics are important, but their skills are also relevant to understand their practices. Their IQ results, their inhibitory control and attention, and some personality trait as measured by the Big Five are important and have a strong relationship with CLASS results.

References

Araujo, M.C., Carneiro, P., Cruz-Aguayo, Y. & Schady, N. (2016). Teacher Quality and Learning Outcomes in Kindergarten. *Quarterly Journal of Economics*, 131(3), 1415-1453.

Blair, C. (2002). School Readiness: Integrating Cognition and Emotion in a Neurobiological Conceptualization of Children's Functioning at School Entry. *American Psychologist*, 57(2), 111-127.

Burchinal, M., Campbell, F., Bryant, D., Wasik, B., & Ramey, C. (1997). Early Intervention and Mediating Processes in Cognitive Performance of Children of Low-Income African American Families. *Child Development*, 68(5), 935-954.

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting Child Outcomes at the End of Kindergarten from the Quality of Pre-Kindergarten Teacher–Child Interactions and Instruction. *Applied Development Science*, 12(3), 140-153.

Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold Analysis of Association between Child Care Quality and Child Outcomes for Low-Income Children in Pre-Kindergarten Programs. *Early Childhood Research Quarterly*, 25(2), 166-176.

Cadima, J., Leal, T., & Burchinal, M. (2010). The Quality of Teacher–Student Interactions: Associations with First Graders' Academic and Behavioral Outcomes. *Journal of School Psychology*, 48(6), 457-482.

Campbell, F., & Ramey, C. (1994). Effects of Early Intervention on Intellectual and Academic Achievement: A Follow-up Study of Children from Low-Income Families. *Child Development*, 65(2), 684-698.

Campbell, F., & Ramey, C. (1995). Cognitive and School Outcomes for High-Risk African-American Students at Middle Adolescence: Positive Effects of Early Intervention. *American Educational Research Journal*, 32(4), 743-772.

Connor, C., Morrison, F., & Petrella, J. (2004). Effective Reading Comprehension Instruction: Examining Child x Instruction Interactions. *Journal of Educational Psychology*, 96(4), 682-698.

Denham, S., & Weissberg, R. (2004). Social-Emotional Learning in Early Childhood: What we Know and Where to go from Here. In E. Chesebrough, P. King, T. Gullota & M. Bloom (Eds.). *A Blueprint for the Promotion of Pro-Social Behavior in Early Childhood* (pp. 13-50). New York: Kluwer Academic/Plenum Publishers.

Downer, J.T., Booren, L., Lima, O., Luckner, A. & Pianta, R. (2010) The Individualized Classroom Assessment Scoring System (inCLASS): Preliminary Reliability and Validity of a System for Observing Preschoolers' Competence in Classroom Interactions. *Early Childhood Research Quarterly*, 25(1), 1-16.

Fox, S., Levitt, P., & Nelson, C. III (2010). How the Timing and Quality of Early Experiences Influence the Development of Brain Architecture. *Child development*, 81(1), 28-40.

Greenberg, M., Domitrovich, C., & Bumbarger, B. (2001). The Prevention of Mental Disorders in School-Aged Children: Current State of the Field. *Prevention & Treatment*, 4(1), 1-59.

Halpern, R. (1986). Effects of Early Childhood Intervention on Primary School Progress in Latin America. *Comparative Education Review*, 30(2), 193-215.

Hamre, B., & Pianta, R. (2005). Can Instructional and Emotional Support in the First-Grade Classroom Make a Difference for Children at Risk of School Failure? *Child Development*, 76(5), 949-967.

Hamre B., Pianta, R., Burchinal, M., Field, S., LoCasale-Crouchs, J., Downer, J.T., Howes, C., La Paro, K., & Scott-Little, C. (2011). A Course on Effective Teacher-Child Interactions. NCRECE Research Brief, retrieved from <https://curry.virginia.edu/ncrece-research-briefs>

Hanushek, E. (2003). The Failure of Input-Based Schooling Policies. *Economic Journal*, 113(485), 64-98.

Howes, C., & Hamilton, C. (1993). The Changing Experience of Child Care: Changes in Teachers and in Teacher-Child Relationships and Children's Social Competence with Peers. *Early Childhood Research Quarterly*, 8(1), 15-32.

Hu, B.Y., Fan, X., Gu, C., & Yang, N. (2016). Applicability of the Classroom Assessment Scoring System in Chinese Preschools based on Psychometric Evidence. *Early Education and Development*, 27(5), 714-734.

Kontos, S. and Wilcox-Herzog, A. (1997) Teachers' Interactions with Children: Why are they so Important?. *Young Children*, 52(2), 4-12

Lazar, I., Darlington, R., Murray, H., Royce, J., Snipper, A., & Ramey, C. (1982). Lasting Effects of Early Education: A Report from the Consortium for Longitudinal Studies. *Monographs of the Society for Research in Child Development*, 47(2-3), 1-151.

La Paro, K., Hamre, B., Locasale-Crouch, J., Pianta, R., Bryant, D., Early, D., Clifford, R., Barbarin, O., Howes, C., & Burchinal, M. (2009). Quality in Kindergarten Classrooms: Observational Evidence for the Need to Increase Children's Learning Opportunities in Early Education Classrooms. *Early Education and Development*, 20(4), 657-692.

La Paro, K., Pianta, R., & Stuhlman, M. (2004). The Classroom Assessment Scoring System: Findings from the Prekindergarten Year. *The Elementary School Journal*, 104(5), 409-426.

Leyva, D., Weiland, C., Barata, M., Yoshikawa, H., Snow, C., Treviño, E., & Rolla, A. (2015). Teacher-Child Interactions in Chile and Their Associations with Pre-Kindergarten Outcomes. *Child Development*, 86(3), 781-799.

Mashburn A.J., Pianta R.C., Hamre B.K., Downer J.T., Barbarin O.A., Bryant D., Burchinal, M., Early, D.M., & Howes, C. (2008). Measures of Classroom Quality in Pre-Kindergarten and Children's Development of Academic, Language, and Social Skills. *Child Development*, 79(3), 732-749.

Merritt, E., Wanless, S., Rimm-Kaufman, S., Cameron, C., Peugh, J. (2012). The Contribution of Teachers' Emotional Support to Children's Social Behaviors and Self-Regulatory Skills in First Grade. *School Psychology Review*, 41(2), 141-159.

National Institute of Child Health and Human Development (NICHD) & Early Child Care Research Network. (2002). The Relation of Global First-Grade Classroom Environment to Structural Classroom Features and Teacher and Student Behaviors. *The Elementary School Journal*, 102(5), 367-387.

Peisner-Feinberg, E., Burchinal, M., Clifford, R., Culkin, M., Howes, C., Kagan, S., Yazejian, N. (2001). The Relation of Preschool Child-Care Quality to Children's Cognitive and Social Developmental Trajectories through Second Grade. *Child Development*, 72(5), 1534-1553.

Perry, K., Donohue, K., & Weinstein, R. (2007). Teaching Practices and the Promotion of Achievement and Adjustment in First Grade. *Journal of School Psychology, 45*(3), 269-292.

Pianta, R., La Paro, K. & Hamre, B. (2008). *Classroom Assessment Scoring System – CLASS*. Baltimore, MD: Brookes.

Pianta, R., & Nimetz, S. (1991). Relationships between Children and Teachers: Associations with Classroom and Home Behavior. *Journal of Applied Developmental Psychology, 12*(3), 379-393.

Phillips, D., Mekos, D., Scarr, S., McCartney, K. and Abbott-Shim, M. (2000). Within and Beyond the Classroom Door: Assessing Quality in Child Care Centers. *Early Childhood Research Quarterly, 15*(4), 475-496.

Ponitz, C., Rimm-Kaufman, S., Brock, L., & Nathason, L. (2009). Early Adjustment, Gender Differences, and Classroom Organizational Climate in First Grade. *The Elementary School Journal, 110*(2), 142-162.

Raver, C. (2004). Placing Emotional Self-Regulation in Sociocultural and Socioeconomic Contexts. *Child Development, 75*(2), 346-353.

Rimm-Kaufman, S., Curby, T., Grimm, K., Nathanson, L., & Brock, L. (2009). The Contribution of Children's Self-Regulation and Classroom Quality to Children's Adaptive Behaviors in the Kindergarten Classroom. *Developmental Psychology, 45*(4), 958-972.

Rudasill, K., Gallagher, K., & White, J. (2010). Temperamental Attention and Activity, Classroom Emotional Support, and Academic Achievement in Third Grade. *Journal of School Psychology, 48*(2), 113-134.

Schweinhart, L., Montie, J., Xiang, Z., Barnett, S., Belfield, C., & Nores, M. (2005). *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40*. Ypsilanti, MI: High/Scope Press.

Taylor, C., Ishimine, K., Cloney, D., Cleveland, G., & Thorpe, K. (2013). The Quality of Early Childhood Education and Care Services in Australia. *Australasian Journal of Early Childhood, 38*(2), 13-21.

Figure 1: CLASS Domains and Dimensions

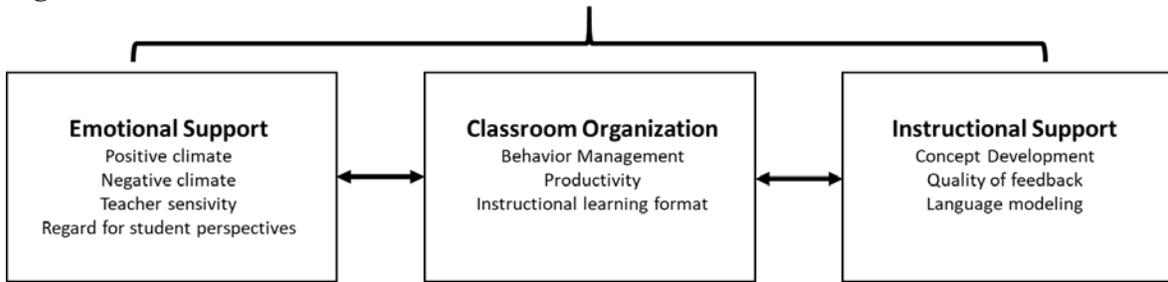
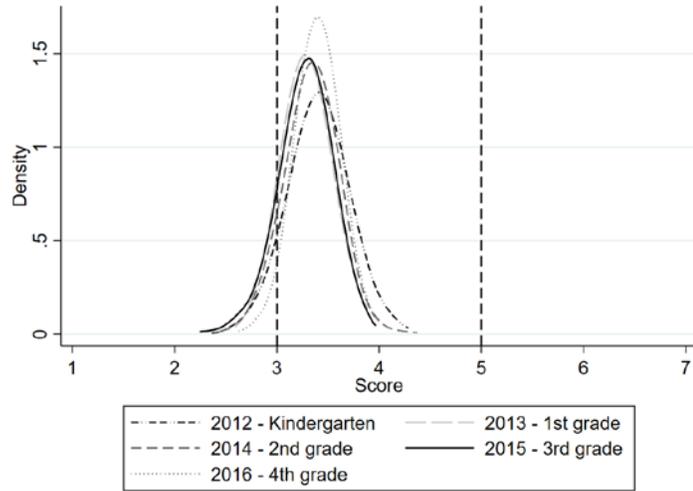
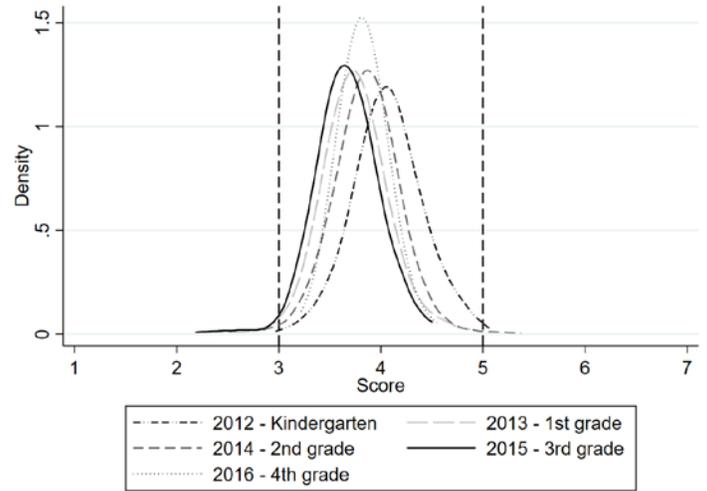


Figure 2: CLASS score by year

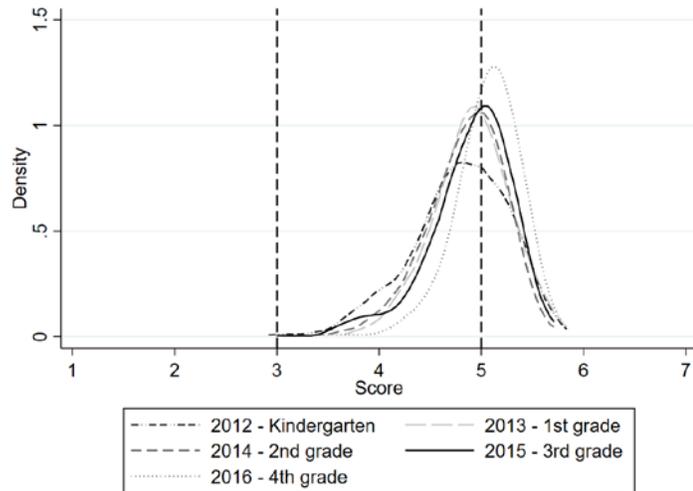
Panel A: Total Score



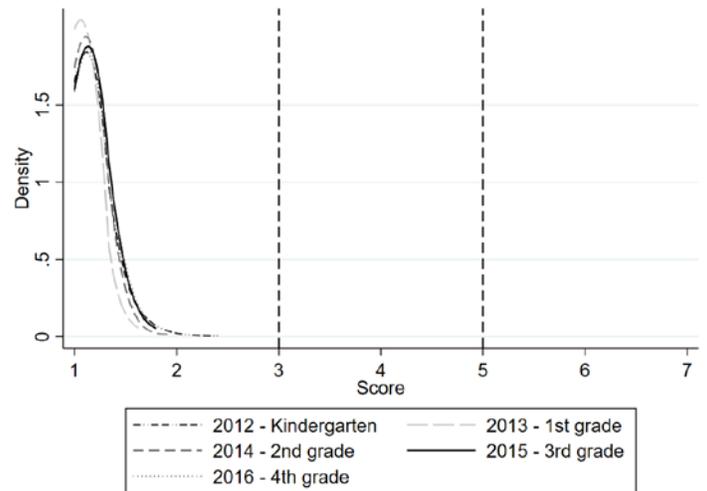
Panel B: Socioemotional Support



Panel C: Classroom Organization

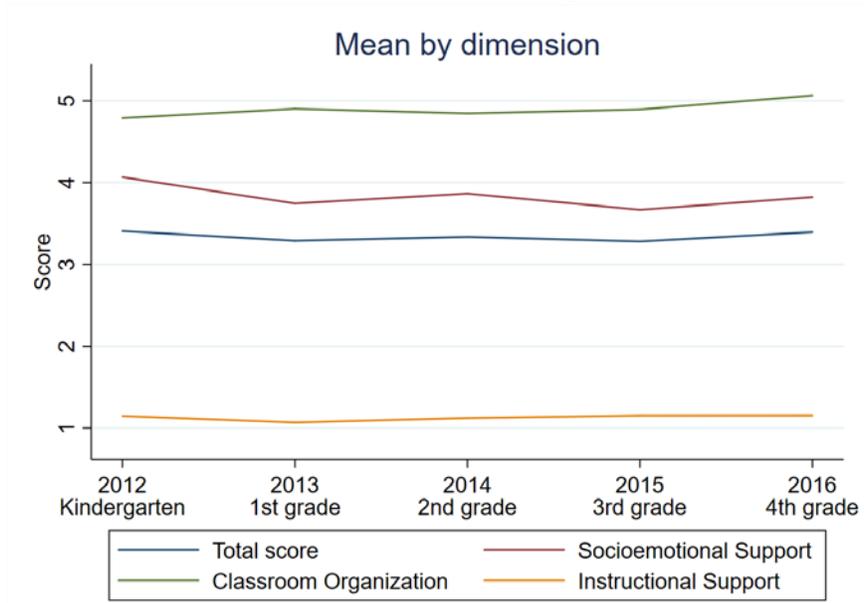


Panel D: Instructional Support



Note: The figure graphs univariate densities of the CLASS score of Kindergarten, 1st grade, 2nd grade, 3rd grade and 4th grade. The CLASS is scored on a 1-7 scale, scores of 1-2 indicate low quality, scores of 3-5 indicate medium quality, and scores 6-7 represent high quality. Calculations are based on an Epanechnikov kernel with optimal bandwidth.

Figure 3: Mean of CLASS (total and by dimension) by grade



Note: The figure shows the mean of the total CLASS and each domain by grade (Kindergarten, 1st grade, 2nd grade, 3rd grade and 4th grade). The CLASS is scored on a 1-7 scale, scores of 1-2 indicate low quality, scores of 3-5 indicate medium quality, and scores 6-7 represent high quality.

Table 1. Summary Statistics

	Mean	S.D.	Obs.
Age	44.27	10.16	2,295
Proportion female	0.87	0.34	2,314
Proportion tenured	0.78	0.41	2,302
Experience (Time working as a teacher)	17.89	10.46	2,308

Table 2. Effect of teacher's characteristics on CLASS score (Panel of teachers)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<u>Total Score</u>							
Tenured teacher	0.376*** (0.056)				0.330*** (0.056)	0.288*** (0.057)	0.348*** (0.056)
Time teaching at the school		0.018*** (0.003)			0.016*** (0.003)		
Time working as a teacher			0.025*** (0.003)			0.020*** (0.003)	
Teacher has 4 years of experience or more				0.400*** (0.095)			0.310*** (0.096)
<u>Emotional Support Score</u>							
Tenured teacher	0.373*** (0.055)				0.336*** (0.056)	0.297*** (0.057)	0.346*** (0.056)
Time teaching at the school		0.015*** (0.003)			0.013*** (0.003)		
Time working as a teacher			0.022*** (0.003)			0.017*** (0.003)	
Teacher has 4 years of experience or more				0.402*** (0.095)			0.306*** (0.096)
<u>Classroom Organization Score</u>							
Tenured teacher	0.316*** (0.056)				0.268*** (0.056)	0.231*** (0.057)	0.289*** (0.056)
Time teaching at the school		0.019*** (0.003)			0.016*** (0.003)		
Time working as a teacher			0.023*** (0.003)			0.019*** (0.003)	
Teacher has 4 years of experience or more				0.380*** (0.095)			0.310*** (0.096)
<u>Instructional Support Score</u>							
Tenured teacher	0.195*** (0.056)				0.174*** (0.056)	0.153*** (0.058)	0.193*** (0.057)
Time teaching at the school		0.009*** (0.003)			0.007*** (0.003)		
Time working as a teacher			0.012*** (0.003)			0.010*** (0.003)	
Teacher has 4 years of experience or more				0.070 (0.095)			0.030 (0.097)
Observations	2,289	2,295	2,295	2,295	2,289	2,289	2,289

Note: All regressions include teacher's age in years, its square and gender, and a year fixed effect. The CLASS is scored on a 1-7 scale, scores of 1-2 indicate low quality, scores of 3-5 indicate medium quality, and scores 6-7 represent high quality.

Table 3. Effect of Kindergarten teacher's tests on CLASS score

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	<u>Panel A: Total Score</u>				<u>Panel C: Classroom Organization Score</u>			
Neuroticism	-0.139*** (0.051)			-0.139*** (0.050)	-0.090* (0.051)			-0.089* (0.051)
Extraversion	0.078 (0.053)			0.073 (0.052)	0.091* (0.053)			0.088* (0.052)
Openness	0.035 (0.054)			-0.001 (0.055)	0.063 (0.054)			0.035 (0.055)
Agreeableness	0.094* (0.054)			0.054 (0.055)	0.061 (0.054)			0.029 (0.055)
Conscientiousness	-0.055 (0.055)			-0.057 (0.055)	-0.026 (0.056)			-0.029 (0.055)
Inhibitory control & attention		0.133*** (0.049)		0.066 (0.052)		0.122** (0.049)		0.068 (0.052)
IQ			0.199*** (0.047)	0.145*** (0.055)			0.172*** (0.047)	0.110** (0.055)
	<u>Panel B: Emotional Support Score</u>				<u>Panel D: Instructional Support Score</u>			
Neuroticism	-0.192*** (0.051)			-0.193*** (0.050)	-0.025 (0.051)			-0.024 (0.051)
Extraversion	0.082 (0.052)			0.075 (0.052)	-0.027 (0.053)			-0.029 (0.053)
Openness	-0.036 (0.054)			-0.074 (0.055)	0.105* (0.055)			0.082 (0.056)
Agreeableness	0.097* (0.054)			0.056 (0.055)	0.095* (0.054)			0.069 (0.056)
Conscientiousness	-0.072 (0.055)			-0.074 (0.055)	-0.041 (0.056)			-0.043 (0.056)
Inhibitory control & attention		0.106** (0.049)		0.038 (0.052)		0.119** (0.049)		0.074 (0.053)
IQ			0.184*** (0.048)	0.160*** (0.055)			0.143*** (0.048)	0.080 (0.056)

Note: All regressions include teacher's age in years, its square and gender. The CLASS is scored on a 1-7 scale, scores of 1-2 indicate low quality, scores of 3-5 indicate medium quality, and scores 6-7 represent high quality.