University of Zurich[UZH]

URPP Equality of Opportunity

TeacherEffects on Student Well-Being and Mental Health

Ana Brás-Monteiro
Marius Opstrup-Morthorst

Equality of Opportunity Research Series #92
September 2026

University of Zurich[UZH]

URPP Equality of Opportunity

URPP Equality of Opportunity Discussion Paper Series No. 92, September 2026

TeacherEffects on Student Well-Being and Mental Health

Ana Brás-Monteiro
University of Zurich
ana.brasmonteiro@econ.uzh.ch

Marius Opstrup-Morthorst
Aarhus University
mmorthorst@econ.au.dk

URPP Equality of Opportunity, University of Zurich, Schoenberggasse 1, 8001 Zurich, Switzerland
info@equality.uzh.ch, www.urpp-equality.uzh.ch

Teacher Effects on Student Well-Being and Mental Health

Ana Brás-Monteiro, Marius Opstrup Morthorst*

28 August 2026

Abstract

Mental health challenges among children and adolescents are a growing concern, with substantial consequences for educational attainment and adult labor market outcomes. This paper studies the role of teachers in influencing student well-being and subsequent mental health outcomes. Using linked administrative and survey data from Denmark, we construct a value-added measure of teachers' effectiveness in improving student well-being during middle school and examine its longer-run effects. We find that teachers have meaningful impacts on students' contemporaneous well-being. These effects persist several years later: students exposed to teachers with higher well-being value-added during middle school report higher well-being in high school. Going beyond survey-based measures, we find persistent reductions in mental health care utilization and depression diagnoses in adolescence. We further show that our measure of well-being value-added is only weakly correlated with test score value-added, and that the two dimensions shape distinct long-run outcomes: well-being value-added primarily reduces mental health care utilization and depression diagnoses, while test score value-added reduces anxiety diagnoses and has stronger effects on academic outcomes.

*Brás-Monteiro: University of Zürich (ana.brasmonteiro@econ.uzh.ch). Morthorst: Aarhus University (mmorthorst@econ.au.dk). Ana Brás-Monteiro is very grateful to Ulf Zölitz, Ana Costa-Ramón, and Jonah Rockoff for their guidance and support. We gratefully acknowledge support from the UZH University Research Priority Program "Equality of Opportunity".

1 Introduction

Mental health problems among children and adolescents impose large public health and economic costs. They are among the leading causes of illness and disability during adolescence worldwide (WHO, 2025).[1] These problems have consequences that extend well beyond childhood, predicting lower educational attainment, employment, and earnings later in life (Currie, 2024). What determines children's well-being and mental health?

Teachers may play a key role. Children spend more than 30 hours a week in school, an environment that shapes their well-being and mental health (Jackson et al., 2020; Hansen, Sabia and Schaller, 2024), and outside their own families, teachers are the adults with whom they interact most throughout childhood. While a large literature documents teachers' influence on student test scores (Koedel, Mihaly and Rockoff, 2015; Chetty, Friedman and Rockoff, 2014*a*), it is increasingly recognized that teachers shape multiple dimensions of students' development (Jackson, 2018; Petek and Pope, 2023). Yet little is known about whether and how teachers influence students' well-being and mental health.

This paper provides the first evidence that teachers shape students' well-being and mental health. We combine rich, population-wide education and health registers with repeated student well-being surveys in Denmark to construct a measure of teachers' effectiveness in improving student well-being — teachers' well-being value-added — and examine its consequences for students' mental health outcomes during adolescence. We also relate this measure to test-score-based measures of teacher quality to study whether the two dimensions capture distinct aspects of teacher effectiveness and affect distinct long-run outcomes. Understanding this is crucial for designing teacher hiring and evaluation policies that identify teachers who can improve not only academic achievement but also students' well-being and mental health.

We start our analysis by documenting a strong relationship between student well-being in middle school and a broad set of later outcomes. Students with higher well-being go on to perform better academically in high school and are substantially less likely to use mental health care services or receive psychiatric diagnoses during adolescence, even after conditioning on test scores.

Motivated by this relationship, we construct a value-added measure of teachers' effectiveness in improving student well-being during middle school. Following Chetty, Friedman and Rockoff (2014*a*), we predict each teacher's effect in a given year from the well-being of the students she taught in all other years, controlling for lagged well-being, lagged test scores,

[1] Globally, one in seven children worldwide experiences a mental health condition. In Denmark, the setting of this paper, roughly 15% of children have been diagnosed with a mental illness by age 18 (Dalsgaard et al., 2020).

and student, classroom, and school characteristics, and we allow teacher quality to drift over time. We find substantial and meaningful variation in teachers' impacts on student well-being, comparable in magnitude to teacher effects on other non–test score outcomes documented in the literature. A one–standard deviation increase in teacher well-being value-added raises student well-being by 0.11 standard deviations.

Observable teacher characteristics, however, explain no more than 1% of the variation in well-being value-added. These results are broadly consistent with the test score value-added literature (Rockoff, 2004; Bau and Das, 2020). Nonetheless, we find that female teachers and more experienced teachers are associated with modestly higher well-being value-added. Female teachers score modestly higher, and well-being value-added increases monotonically with experience, with a gap of 0.09 standard deviations between the most and least experienced. At the same time, students themselves appear to recognize effective well-being teachers: using our leave-year-out measure of well-being value-added, we find that a one standard deviation increase raises students' ratings of teacher quality by 0.104 standard deviations.

We then examine how exposure to teachers with higher value-added affects students' outcomes several years later, including self-reported well-being in high school, mental health care utilization and psychiatric diagnoses between ages 16 and 18, and academic performance in high school. Importantly, we use students' self-reports to construct teacher well-being VA, but our longer-term outcomes are based on administrative measures, such as visits and diagnoses. We find that students exposed to teachers with higher well-being value-added report significantly higher levels of well-being during high school. Moreover, they are substantially less likely to use mental health care services between ages 16 and 18, and they experience reductions in depression diagnoses. Taken together, the results suggest that teachers' impacts on well-being extend beyond subjective survey responses to longer-run, objective mental health outcomes.

We next ask whether these effects are evenly distributed across students, or whether certain groups benefit disproportionately. Effects on self-reported high school well-being are homogeneous across subgroups. Effects on mental health outcomes, however, are concentrated among students with higher baseline achievement and well-being. These patterns suggest that while teachers improve the subjective well-being of all students, the downstream protection against mental health risk is concentrated among those who were already doing relatively well.

Beyond heterogeneity across students, we also find meaningful variation in effects across the grade of exposure to the teacher. We find that the effects of well-being value-added are present across all grades of middle school, but are especially pronounced at grades 4 and 7, which mark educational transitions in Denmark, each associated with changes in curriculum and instruction. Effects on high school well-being and mental health care utilization are larger in these transition grades than in others, suggesting that exposure to high well-being VA teachers is especially consequential when students face new challenges in the school environment.

Finally, we also examine how teachers' well-being value-added relates to their test score value-added. We show that well-being value-added captures a distinct dimension of teacher quality: it is only weakly correlated with test score value-added, with a correlation of approximately 0.25. Consistent with the multidimensional view of teacher quality, we find that test score value-added has much stronger effects on long-run academic outcomes, whereas well-being value-added has the strongest effects on students' long-run well-being and mental health care utilization. When it comes to specific diagnoses, well-being value-added reduces depression while test score value-added reduces anxiety. Disaggregating high school well-being into its sub-dimensions helps explain this contrast: we find that being exposed to a high well-being VA teacher during middle school improves social dimensions of school life in high school, which may be more closely linked to depression risk, while test score VA improves academic well-being but has no positive effect on social or emotional dimensions, consistent with the idea that academic pressure may be more relevant for anxiety. Most of these patterns persist when both dimensions are included jointly in a predictive regression, indicating that neither measure simply proxies for the other.

To address concerns of sorting and selection biases, all models condition on a rich set of lagged outcomes and student, classroom, and school characteristics. We further implement a series of empirical tests to assess whether the estimated relationships can be interpreted causally. Following Chetty, Friedman and Rockoff (2014*a*,*b*), we examine forecast bias, test for selection on observable characteristics excluded from the value-added model, and exploit quasi-experimental variation arising from changes in teaching staff across cohorts. Across these exercises, we find little evidence of systematic sorting that would materially bias either the contemporaneous or long-run estimates of teachers' impacts on student well-being and mental health.

Our paper contributes to two different strands of literature. First, we contribute to the growing literature on child and adolescent mental health by focusing on the role of the school environment, and teachers in particular, in shaping student well-being and, consequently, their mental health. A large body of work shows that mental health is shaped by early-life conditions, including family background and early shocks, as well as post-natal environments such as income shocks and parental job loss (Bütikofer et al., 2023; Adhvaryu, Fenske and Nyshadham, 2019; Persson and Rossin-Slater, 2018; Baird, De Hoop and Özler, 2013; Schaller and Zerpa, 2019). More recent studies emphasize other aspects of schools, including the role of school starting age, instructional intensity, and school violence (Dee and Sievertsen, 2018; Marcus et al., 2020; Rossin-Slater et al., 2020). Relatively little is known about the role of individual teachers in shaping student well-being and mental health. Our paper addresses this gap by documenting that teachers can have meaningful positive impacts on student well-being that lead to persistent improvements in mental health outcomes.

Our paper also contributes to the literature on teacher value-added and the multidimensional nature of teacher quality. Chetty, Friedman and Rockoff (2014*a*,*b*) document substantial heterogeneity in teachers' impacts on test scores and show that exposure to high–value-added teachers has lasting effects on educational attainment and earnings. A growing body of work further demonstrates that teacher quality extends beyond test scores, with teachers affecting absences, suspensions, grades, socio-emotional competencies, and even criminal justice contact (Jackson, 2018; Petek and Pope, 2023; Rose, Schellenberg and Shem-Tov, 2022; Backes et al., 2024). While this literature makes a compelling case that teachers shape non-cognitive skills, these skills are typically inferred from observed behaviors and school records rather than measured directly. We extend this literature by bringing direct measures of student well-being and population-wide mental health outcomes into the analysis of teacher effectiveness. Using the value-added framework, we show that teachers meaningfully influence student well-being, and that this dimension of teacher quality is only weakly correlated with test score value-added. We also document that this new measure of teacher quality has lasting effects on students' mental health outcomes. These findings highlight the non-academic channels through which teachers shape students' long-run trajectories.

Closest to our work, a handful of recent studies use survey-based or standardized assessments to estimate teachers' effects on students' socio-emotional skills and personality traits (Kraft, 2019; Fleche, 2017; Montonen, 2025). In contrast, we focus on student well-being and, to our knowledge, are the first to link teacher value-added to population-wide administrative measures of mental health care utilization and psychiatric diagnoses.

The remainder of the paper is organized as follows. Section 2 describes the institutional setting and data, covering the Danish education and health care systems as well as the construction of the analysis sample and key variables. Section 3 outlines the empirical method for estimating teacher value-added and its long-run effects. Section 4 presents the main results on the effects of teachers' well-being value-added on students' mental health and academic outcomes. Section 5 assesses the validity of the value-added estimates and the long-run effects. Section 6 studies the multidimensional nature of teacher quality, documenting the within-teacher correlation between well-being and test score value-added and showing that each dimension independently affects distinct long-run outcomes. Section 7 concludes.

2 Setting and Data

2.1 The Education System

Denmark offers a high-quality, publicly funded educational system. Compulsory schooling (*Grundskole*) spans ten years, beginning in the calendar year when children turn six, and is organized as an integrated school covering both primary and lower secondary education. It consists of one pre-school year (Grade 0), nine years of primary and lower secondary education (Grades 1–9), and an optional tenth grade.[2] While compulsory education is primarily publicly funded and free of charge (as is almost all post-compulsory education, including higher education institutions), approximately 18 percent of students attend private schools, where they pay tuition fees.

Upon entering Grade 0, children are grouped into classes and typically remain in the same classroom, with an average class size of 21 pupils, until Grade 9.[3] While schools have some discretion over class formation, national guidelines set out in the *Folkeskole Act* state that, as the Folkeskole is not an examination-oriented school, the main rule is that pupils should attend classes with peers of the same age rather than being grouped by subject-specific proficiency.

Classes are taught by multiple subject-specific teachers. However, each class also has a designated class teacher with primary responsibility for monitoring students' academic and social development, as well as ensuring coordination with other teachers and communication with families. Usually, the Danish teacher is the class teacher.

The schools administer several standardized tests and examinations during compulsory schooling. The national test program, introduced in 2010, includes ten standardized tests administered in Grades 2 through 8, covering reading, mathematics, and other subjects (Beuchert and Nandrup, 2018). Compulsory education then concludes with five mandatory exams in Grade 9. The results of these exams contribute to the student's final GPA. Additionally, students receive year marks from their subject-specific teacher for their overall performance

[2]Throughout this paper, we refer to Grades 4–9 as "middle school", a term broadly equivalent to the upper-primary and lower-secondary years in other countries, including the United States.

[3]However, some schools may reshuffle children in Grade 3 and/or Grade 7 to try new group dynamics.

throughout the school year, which they use for applying to further education. Schools also administer yearly well-being surveys; further details are provided in the following subsection.

After Grade 9, students choose between several post-compulsory paths. Approximately 36% enroll directly in academic high school, around 51% first complete an optional additional year of schooling known as 10th grade, and 6% enter VET programs directly.[4] The large majority of students who complete 10th grade subsequently enroll in academic high school. Academic high school programs prepare students primarily for higher education, whereas VET programs equip students with skills for a specific trade or industry, with limited opportunities for direct enrollment into higher education. Students may freely enroll in academic high school if their GPA is at least 5, or in the two-year preparatory program if their GPA is at least 4, on the Danish 7-point grading scale ranging from −3 to 12.[5] Within academic high school, students choose among several tracks that differ in curricular focus, including a general program (STX), a business-oriented program (HHX), and a technical program (HTX).

2.2 The Health Care System

Denmark operates a universal, publicly funded health care system in which general practitioners (GPs) serve as the primary point of contact and act as gatekeepers to specialist care. Health care is provided free of charge at the point of use, with specialist visits fully covered under the national health insurance scheme provided the patient holds a GP referral.

For children and adolescents with mental health concerns, the system follows a stepped model. GPs provide initial assessment and can prescribe psychotherapy, but guidelines issued by the Danish Health and Medicines Authority recommend that children at risk of a mental health disorder be referred to specialist care: mild cases to private practice child psychiatrists, and more severe cases to hospital-based psychiatric departments. A key feature of the system is that prescription drugs for mental health conditions for individuals below age 18 must be initiated by a specialist in psychiatry rather than a GP. As a result, all children receiving pharmacological treatment for conditions such as depression or ADHD will have had formal contact with a specialist.

[4] Based on the 2018 cohort; (Statistics Denmark, 2024).

[5] As of 2018, students must also pass all mandatory exams. Below these thresholds, students may still enroll in VET programs (which require a minimum grade of 02 in Danish and mathematics) or apply to academic programs through individual assessment by the receiving school.

2.3 Data

Our analysis draws on several Danish administrative registers that can be linked using unique individual identifiers. We begin with the Education Register, which contains enrollment records for all students attending public schools in Denmark. Our sample includes students enrolled in Grades 4 through 9 during the 2015/2016–2019/2020 academic years. We link these data to the Teacher Register to assign students to their subject-specific teachers. The analysis uses teacher-student links in Danish and Mathematics, as these are the two subjects for which we have frequent information on test scores. We further link our sample to the Central Population Register to obtain demographic characteristics and parental identifiers. Below, we describe the key variables used in the analysis and the additional data sources from which they are drawn.

Well-Being. To measure well-being, we use the responses to the nationwide well-being survey, which has been conducted annually since 2015 for all students in compulsory education between January and April of every year. The survey is administered electronically, with a designated teacher in each classroom overseeing the process. The teacher further emphasizes the importance of honest responses and assures students that their individual answers will remain confidential and will not be shared with parents, teachers, or other school personnel. Even though it is mandatory for schools to administer the survey, it is not mandatory for students to complete it. Regardless, the response rate is very high, roughly 85%, and highest for earlier grades.

The questionnaire contains 40 questions addressing different aspects of well-being and school life in general. Among the full list of questions, we use the 18 questions that enter into the academic and social well-being subscales, as defined by Danish Ministry of Education (2016). The translated version of the questions can be found in Appendix B.1. The answers are coded to range from one to five, with higher values indicating more positive well-being. We construct the final well-being measure by computing the within-individual average of the answers provided and standardizing it by cohort. The distribution of this score, before standardization, is depicted in Appendix Figure A.1.

Test scores. For grades 2 through 8, we use the results of the national tests in Danish and Mathematics. Specifically, we use reading test scores in Grades 2, 4, 6, and 8, and mathematics test scores in Grades 3, 6, and 8 (for 2018 only). All test scores are standardized by subject and cohort. For Grade 9, we use results from the exit exams, which we obtain from administrative records on Grade 9 exit exams. We focus on final scores in Danish and Mathematics, which we standardize by cohort.

Education outcomes. We study two post-compulsory education outcomes: enrollment in academic high school and final high school GPA. For the former, we define the variable to take value 1 if the student has ever enrolled in an academic high school program. For the latter, we use administrative records on high school graduation, restricted to academic high school graduates, which we standardize by graduation year and academic track.[6]

High School Well-Being. To measure well-being in high school, we draw on data from the Danish Well-Being Survey administered in academic high school programs. Although the questionnaire differs from that used during compulsory schooling, Danish Ministry of Education (2023) constructs the same two domains of academic and social well-being. We construct the well-being measure analogously and focus on the first year of high school. As supplementary analysis, we also examine the three additional sub-indices constructed by the Ministry of Education: Learning Environment, Pressure & Concerns, and Bullying.[7] Participation in this survey is voluntary, resulting in a smaller sample. Note that, unlike the mental health outcomes which cover all students, this measure is by construction limited to the subsample who enrolled in academic high school.

Mental health outcomes. The mental health outcomes capture health care utilization and diagnosis. We measure mental health care utilization using the *Psychiatric Central Research Register* and the *Health Insurance Register*. The former includes all inpatient, outpatient and ER visits to psychiatric departments in public and private hospitals. The latter captures reimbursements to private clinics (GPs and specialists) for services covered by national health insurance. We use the same dataset to identify receipt of psychological counseling.

Among hospital-based visits, we define a psychiatric diagnosis using ICD-10 codes starting with "F". We further define indicators for different types of diagnosis, such as depression, anxiety, behavioral and emotional disorders and ADHD.[8] In our main results, we focus on mental health care utilization between the ages of 16 and 18.

Analysis sample. We treat the student-subject-year as the relevant unit of observation. Therefore, each observation includes a student's subject-specific test score for a given year, their lagged test score, their contemporaneous well-being measure, lagged well-being, and the long-run outcomes.

To ensure sufficient information to estimate teacher effects, we drop teacher–subject–year cells with fewer than ten students, corresponding approximately to the bottom one percent of

[6] Academic high school programs differ in curriculum and do not all include the same coursework, making raw GPA scores not directly comparable across tracks.

[7] See Appendix B.3 for the translated version of the questions entering each sub-index.

[8] See Appendix B.2 for details on the exact codes used to construct the mental health outcomes.

the student count distribution. We further impose restrictions based on the availability of well-being data, as this is the basis of our teacher quality measure. Specifically, we drop classrooms in which fewer than 20 percent of students have non-missing values for both current and lagged well-being. Finally, we restrict the analysis to student–subject–year observations with complete information on current and lagged well-being.

Descriptive statistics. Table 1 provides summary statistics for the full sample and for the final analysis sample, after all restrictions described above. As displayed, we see that the restricted sample is quite similar to the full sample in terms of observable characteristics and outcomes.

2.4 Student Well-Being and Later Outcomes

We begin by documenting the relationship between student well-being in middle school and a range of later outcomes. Table 2 reports regressions of high school academic outcomes, adolescent well-being, mental health care utilization, and psychiatric diagnoses during ages 16 to 18, on standardized measures of student well-being measured in middle school, conditional on the same set of controls we use to compute the value-added measure. We find that higher student well-being in middle school is strongly associated with better subsequent well-being in high school and substantially lower mental health care utilization and psychiatric diagnoses during adolescence, even when controlling for test scores. Student well-being is also predictive of later academic outcomes, though test scores are more strongly related to academic performance in high school.

While these relationships are descriptive and may reflect selection on unobservables, they highlight student well-being as a distinct and consequential dimension of student development that is not captured by test scores alone. We view these patterns as a validation of the well-being survey measure, which is particularly important given that it forms the basis of the teacher value-added measure we construct next.

3 Empirical Strategy

3.1 Estimating Teacher Value-Added

We estimate teacher value-added (VA) following the framework developed by Chetty, Friedman and Rockoff (2014*a*), adapting it to measure teachers' impacts on student well-being. The goal is to isolate teachers' persistent contributions to student outcomes by comparing realized outcomes to those predicted from prior outcomes, background characteristics, and classroom context.

Let Y_{ijt} denote the outcome of interest—either test scores or well-being—for student i taught by teacher j in year t. We construct value-added measures by first residualizing the outcome of interest with respect to a rich set of controls X_{ijt} and teacher fixed effects. We flexibly control for prior well-being and test scores using lags of a cubic polynomial of the student's well-being, Danish and math test scores, and we interact these cubics with the student's grade level.[9] We also control for students' ethnicity, gender, age and maternal education. We also include the following class- and school-level controls: (i) cubics in class and school-grade means of prior-year well-being and test scores, each interacted with grade; (ii) class and school-grade means of all the other individual covariates; (iii) class size; and (iv) grade-by-year fixed effects.

Let ν_{ijt} be the residualized outcome. We estimate $\hat{\beta}$ via OLS including teacher fixed effects α_j to prevent teacher quality from being absorbed into the coefficients on student characteristics:

$$Y_{ijt} = \alpha_j + \beta X_{ijt} + \varepsilon_{ijt}. \tag{1}$$

We then form the residual by removing only the estimated effect of observables:

$$\nu_{ijt} = Y_{ijt} - \hat{\beta} X_{ijt} = \mu_{jt} + \epsilon_{it}, \tag{2}$$

where μ_{jt} represents the teacher's contribution to student outcomes in year t and ϵ_{it} captures idiosyncratic student-year shocks. For each teacher-year, we compute the mean residual across students:

$$\bar{\nu}_{jt} = \frac{1}{N_{jt}} \sum_{i \in j,t} \nu_{ijt}, \tag{3}$$

which provides a noisy estimate of the teacher's effectiveness in year t. Because these estimates are subject to sampling error and transitory shocks, we follow Chetty, Friedman and Rockoff (2014*a*) and construct a noise-adjusted measure of persistent teacher quality using a leave-year-out empirical Bayes procedure. Specifically, we predict a teacher's value-added in year t

[9] We exclude observations with missing current or prior well-being in all specifications, as well-being is the primary outcome of interest. In the estimation of well-being value-added, we allow prior test scores in either subject to be missing; when missing, prior test scores are set to zero and flagged with missing indicators interacted with controls for prior well-being. In the estimation of test score value-added, we exclude observations with missing current or prior test scores in the subject for which value-added is being estimated. When prior test scores in the other subject are missing, we set them to zero and include an indicator for missing data interacted with the controls for prior own-subject test scores.

using information from the teacher's residuals in other years:

$$\widehat{\mu}_{jt} = \sum_{\substack{s=t-a \\ s\neq t}}^{t+a} \psi_s \bar{\nu}_{js} \tag{4}$$

where we allow the weights ψ_s to vary by the number of years before and after year t, and a is the maximum number of years available in the data for teacher j. Intuitively, a teacher's performance one year ago might be more predictive than their performance four years ago, because teacher quality can evolve over time, so ψ_{t-1} and ψ_{t+1} might be larger than ψ_{t-4}. By estimating separate weights for each relative year position, the model learns how much each neighboring year's data should contribute. We estimate these weights by minimizing the mean squared error of the forecasts:

$$\boldsymbol{\psi} = \underset{\{\psi_{t-a},\ldots,\psi_{t+a}\}}{\arg\min} \sum_j \left(\bar{\nu}_{jt} - \sum_{\substack{s=t-a \\ s\neq t}}^{t+a} \psi_s \bar{\nu}_{js} \right)^2 \tag{5}$$

This procedure produces leave-year-out jackknife value-added predictions that allow teacher quality to change over time, and it shrinks the value-added predictions to the mean through Bayesian shrinkage.

We implement this methodology separately for test scores and for well-being, yielding two distinct value-added measures. Throughout the paper, all value-added estimates are constructed using leave-year-out predictions, ensuring that contemporaneous outcomes are never mechanically used to construct the teacher quality measures applied to those outcomes. In Section 5, we discuss the identifying assumptions underlying this approach and present a series of tests for forecast bias and sorting, as well as a quasi-experimental analysis exploiting teachers who switch grades and schools.

3.2 Estimating the Long-Term Effects of Teacher Value-Added

Having constructed measures of teacher quality, we examine whether exposure to higher- or lower-quality teachers affects students' medium- and long-run outcomes. We estimate the following specification:

$$y_i = \sum_{k=1}^{K} \delta_k \hat{\mu}_{jkt} + \beta X_{ijt} + \eta_{ijt} \tag{6}$$

where y_i denotes the outcome of interest (e.g., high school GPA or an indicator for mental health care utilization between ages 16 and 18), and $\widehat{\mu}_{jkt}$ represents teacher value-added in dimension k (e.g., test scores or well-being). We normalize each value-added measure by the standard deviation of the corresponding teacher effects, so that δ_k can be interpreted as the effect of a one standard deviation increase in teacher value-added. We estimate models that include each dimension separately as well as specifications that include both jointly. The coefficients δ_k therefore capture how different dimensions of teacher quality predict students' longer-run outcomes.

In order to be able to interpret these estimates as causal, one would need to make sure that unobserved determinants of students' long-term outcomes are uncorrelated with teacher quality, conditional on observables. Given the finding that conditioning on prior scores and other observables is adequate to obtain unbiased estimates of teachers' causal impacts on test scores, one may expect that comparing the long-term outcomes of students assigned to different teachers conditional on the same control vector will yield unbiased estimates of teacher's long-term impacts. However, to rule out some concerns with this assumption, we perform a quasi-experimental analysis that uses teachers switching grades and schools in Section 5 (Chetty, Friedman and Rockoff, 2014*a*).

4 Main Results

4.1 Teacher Value-Added Estimates

We first characterize the magnitude of teacher value-added on student well-being. Table 4 reports the square roots of the estimated variance components, estimated separately for Danish and Math teachers. Row 4 reports the standard deviation of teacher effects, computed from the within-year between-class covariance of mean residuals.

We estimate substantial heterogeneity in teacher well-being effects. The SD of teacher effects on well-being is 0.120 for Danish teachers and 0.101 for Math teachers, with a pooled SD of 0.109. This means, for example, that having a Danish teacher with well-being value-added at the 85th versus 50th percentile would increase student well-being by roughly 0.120 standard deviations. These magnitudes are comparable to benchmark estimates of teacher effects on test scores in prior work (e.g., Chetty, Friedman and Rockoff, 2014*a*; Jackson, 2018; Rose, Schellenberg and Shem-Tov, 2022), suggesting that teachers meaningfully influence students' well-being.

The within-year between-class covariance is identified only for teachers who teach multiple classes in the same year. In smaller schools where a teacher has a single class per year, this moment is unavailable and that teacher does not contribute to the estimation of the within-year between-class covariance. For comparison, Row 5 of Table 4 reports the square root of the lag-1 cross-year covariance, which is available for all teachers with at least two years of data.

The empirical distribution of our jackknife well-being VA estimates is plotted in Figure 1. The standard deviation of these estimates is 0.067, smaller than the true SD reported above, as the best linear predictor shrinks VA estimates toward the sample mean to minimize mean-squared error. For reference, Table 4 and Figure 1 also report the corresponding estimates for test score value-added, which we discuss in more detail in Section 6.

4.2 Teacher Characteristics

Using our well-being VA estimates, we estimate the association between value-added and observed teacher characteristics. Specifically, we relate well-being VA to two sets of characteristics: professional factors, including further education beyond a teaching degree and years of experience, and personal factors, including gender, having children, coming from a teacher family, and having any psychiatric diagnosis. Figure 2 reports the results. Observable characteristics account for less than 1 percent of the variation in well-being VA ($R^2 = 0.005$), consistent with prior evidence that effective teachers are difficult to identify from observable characteristics (Rockoff, 2004). Among the characteristics we examine, years of experience is the strongest predictor: relative to teachers with fewer than six years of experience, teachers with more than 20 years of experience have well-being VA that is approximately 0.09 standard deviations higher, with a monotonically increasing pattern across experience categories. Female teachers are also associated with modestly higher well-being VA, while further education and having a teacher family background show no significant association. Teachers with a psychiatric diagnosis have somewhat lower well-being VA, though the estimate is imprecise.

Despite the limited predictive power of observable characteristics, students themselves appear to recognize effective well-being teachers. We construct a student-perceived teacher quality index from five survey items capturing classroom management and the quality of the teacher-student relationship.[10] Regressing this index on well-being VA — conditional on the

[10]The index is the within-student average of responses to the following questions: (Q18) *If there is noise in the class, teachers can quickly create calm*; (Q23) *Do your teachers show up for class punctually?*; (Q27) *Do your teachers help you learn in ways that work well?*; (Q32) *The teachers are good at supporting me and helping me at school when I need it*; (Q37) *Teachers ensure that students' ideas are used in teaching*. We note that these questions ask students to evaluate their teachers in general, rather than a specific teacher. Since students spend a large fraction of their school day with their Danish and mathematics teachers, we use this index as a proxy for students' perceptions of the teachers in our analysis.

baseline control vector — yields a coefficient of 0.104 (SE = 0.003, $p < 0.001$), indicating that a one standard deviation increase in well-being VA raises student-perceived teacher quality by 0.104 standard deviations.

4.3 Long-Run Effects of Well-Being Value-Added

4.3.1 Well-Being and Mental Health Outcomes

This section tests whether teachers who improve well-being cause improved well-being and mental health outcomes for students in later years. We estimate equation 6 with well-being VA, in which the main outcomes are well-being during high school, whether the student utilized any mental health care services between ages 16 and 18, and whether the student had any psychiatric diagnosis between ages 16 to 18. We present the results using binscatters (see Figure 3) of the relationship between teachers' well-being value-added and each of the outcomes, conditional on the set of controls used to compute value-added measures. The plotted points show the relationship between the mean residualized outcome and the mean residualized value-added variable (with the unconditional mean of the outcome and value-added variable added back in) for 20 equally sized bins of teacher value-added measures. The coefficients and standard errors reported in the figures are from an OLS (ordinary least squares) regression, using the micro data, of the outcome variable on the value-added variable, conditional on the same set of controls.

Figure 3a shows that students with higher well-being VA teachers in middle school report higher levels of well-being during high school. Specifically, a one standard deviation improvement in well-being teacher VA in a single grade increases high school well-being by 0.044 standard deviations.

Appendix Figure A.8 decomposes this effect into its two components — academic and social well-being — and additionally examines the remaining sub-indices constructed by the Ministry of Education. Effects on academic and social well-being are of comparable magnitude, consistent with the headline result. Assignment to a higher well-being VA teacher also improves learning environment by a similar amount, and reduces pressure and concerns and bullying, though these effects are smaller in magnitude.

We next examine whether the improvements in self-reported well-being translate into effects on more objective mental health outcomes, including mental health care utilization and psychiatric diagnoses. Figure 3b shows that assignment to a teacher with higher well-being value-added in a single grade significantly reduces the probability that a student visits a psychologist or psychiatrist between ages 16 and 18. The estimated effect corresponds to a 5.6 percent reduction relative to the mean rate of utilization, indicating that teachers who improve students' well-being also reduce subsequent demand for mental health care. We also find reductions in the likelihood of receiving a psychiatric diagnosis over the same age range, though these effects are smaller and less precisely estimated (Figure 3c).

The presence of consistent effects across both utilization and diagnosis supports the idea that teachers' impacts on well-being extend beyond subjective reports to clinically relevant mental health outcomes. Figure 4 presents estimates separately for different diagnoses. When we disaggregate psychiatric diagnoses by type, the effects are concentrated in depression, with no detectable effects on anxiety.

A natural concern in interpreting mental health care utilization is that it can capture different underlying mechanisms. Higher utilization may reflect worsening mental health, as students seek care in response to distress, but it may also indicate improved access to care or a greater willingness to seek help. While we cannot fully distinguish between these interpretations, the reductions in utilization we observe occur alongside sustained increases in self-reported well-being. Taken together, this pattern is more supportive of improvements in underlying mental health rather than reduced help-seeking.

Heterogeneity. We examine heterogeneity in these effects by re-estimating our main specifications separately for subgroups defined by four baseline characteristics: academic achievement (below/above median prior test scores), well-being (below/above median prior well-being), gender, and mental health risk (below/above median predicted probability of mental health care utilization, based on observable characteristics at baseline). For each split we estimate Equation 6 separately and report the effect of a one standard deviation increase in well-being VA for each subgroup. Figure 5 shows the results across the three main outcomes, together with the p-value from a test of equality of effects across each pair of subgroups.

Effects on self-reported high school well-being are broadly homogeneous, ranging from roughly 0.042 to 0.047 standard deviations across subgroups, with differences that are small and, apart from a modest gender gap, not statistically significant. The effects on mental health care utilization, by contrast, vary systematically with baseline advantage. The reduction is larger for students with high baseline test scores than low (−6.3% vs. −4.0% relative to the group mean, $p < 0.001$) and for students with high prior well-being than low (−7.5% vs. −4.0%, $p = 0.002$). Boys show larger reductions than girls ($p < 0.001$). Differences by predicted mental health risk point the same way, with larger reductions for lower-risk students, though they are not statistically significant ($p = 0.30$).

The pattern is starker for psychiatric diagnoses. Among students with low baseline test scores the estimated effect is essentially zero, whereas among high-achieving students it is a reduction of 4.2% relative to the mean ($p < 0.001$). The same holds for baseline well-being: low-WB students show near-zero effects, while high-WB students show reductions of 5.1% ($p < 0.001$). There is no significant difference by gender ($p = 0.68$), and lower-risk students show somewhat larger reductions than higher-risk students ($p = 0.05$).

Taken together, these results point to a boundary on what a teacher can do alone. Teachers raise the subjective well-being of disadvantaged students as much as that of their peers, but this improvement translates into fewer clinical mental health problems only for students who were already doing comparatively well. For students with lower baseline achievement or well-being, the determinants of clinical risk appear to lie largely outside the classroom, so a supportive year changes how they feel without altering their clinical trajectory. A teacher's protective influence on mental health instead operates among students closer to the margin, where a good year can keep them from crossing a clinical threshold.

Dynamic effects by grade. We next examine whether the effects of well-being value-added vary across grades, following Petek and Pope (2023). Danish compulsory schooling is formally subdivided into three levels: *indskoling* (Grades 1–3), *mellemtrin* (Grades 4–6), and *udskoling* (Grades 7–9), each associated with changes in curriculum and instruction. Grades 4 and 7 therefore mark natural transition points within compulsory schooling, motivating us to examine whether exposure to high well-being VA teachers is especially consequential at these moments. For each grade from 4 to 9, we estimate the effect of a one standard deviation increase in well-being VA on each outcome. Figure 6 reports the results, alongside the sum of coefficients across all grades as an upper bound on the cumulative effect.

Consistent with transitions being particularly consequential, effects on high school well-being are positive in every grade but substantially larger in grades 4 and 7 — approximately 0.08 standard deviations each — compared to the intervening and later grades. The sum of coefficients is 0.26 standard deviations. The effects on mental health care utilization broadly mirror this pattern: reductions are largest in grades 4 and 7, with a near-zero effect in grade 6, and a cumulative sum of −0.051. The effects on psychiatric diagnoses follow the same directional pattern — negative in the earlier grades and fading toward zero by the final grades — but are considerably less precisely estimated, reflecting the lower prevalence and noisier nature of formal diagnoses.

4.3.2 Academic Outcomes

While the previous results show that teachers' well-being value-added has meaningful effects on students' mental health and well-being, it is important to assess whether these effects extend to outcomes such as high school enrollment in academic tracks and grade point average, which are central to educational policy and long-run economic outcomes.

Figures 7 show that teachers' well-being value-added also has positive, though more modest, effects on academic outcomes. Assignment to a teacher with one standard deviation higher well-being value-added increases the probability of enrolling in an academic high school track by 0.007, relative to a mean enrollment rate of 0.69. We also find a positive effect on high school GPA: a one–standard deviation increase in well-being value-added raises students' standardized GPA by 0.007. This effect could be seen as a lower bound on the true effect. If we believe students induced into academic high school by being exposed to higher well-being VA teachers are marginal enrollees, they face a more demanding curriculum than they would have otherwise. This could mechanically attenuate the GPA estimate downward.

Overall, while these magnitudes are smaller than the effects on mental health outcomes, they suggest that teachers who improve student well-being do not do so at the expense of academic success and may modestly improve it. We contrast these effects with those of teachers' test score value-added in Section 6, where we show that test score value-added has substantially larger impacts on academic outcomes but a distinct pattern of effects on mental health.

5 Testing the Identifying Assumptions

5.1 Bias in the VA Estimates

We assess the potential for bias in the teacher value-added estimates using three complementary approaches, following Chetty, Friedman and Rockoff (2014*a*): tests of forecast unbiasedness, analyses of selection on observable characteristics, and a quasi-experimental design exploiting changes in teaching staff across cohorts.

5.1.1 Forecast Bias

First, we test whether the well-being value-added measure is forecast unbiased. Under the stationarity assumption, regressions of contemporaneous residualized outcomes on leave-year-out value-added should yield a coefficient of one by construction. Panel A of Figure 8 shows that this is the case: the estimated relationship is close to one. However, this relationship could be driven by the causal impact of teachers on well-being, or by persistent differences in student characteristics across teachers. Therefore, we focus now on whether there is systematic sorting of students to teachers.

5.1.2 Selection on observables

Although we cannot directly test for sorting on unobserved determinants of student achievement, we can test for sorting on observable characteristics that predict well-being or test score residuals but are excluded from the value-added model. We conduct this test using maternal income and outcome-specific twice-lagged measures — twice-lagged test scores for test score VA and twice-lagged well-being for well-being VA.

We show that, after conditioning on our rich control vector, students from higher income families are largely not sorting to higher well-being VA teachers (Panel B of Figure 8). Using twice-lagged well-being, we estimate a small coefficient of -0.003 (s.e. 0.003) (Panel C of Figure 8), which is close to zero and opposite in sign to prior estimates in the literature.[11] The negative coefficient implies that, if anything, students with higher twice-lagged predicted achievement are weakly sorted toward lower value-added teachers, which would bias our estimates downward rather than upward. In all cases, the magnitude of this relationship is small relative to the relationship between teacher value-added and contemporaneous outcome residuals. Overall, this is consistent with the broader literature on teacher effects, suggesting that VA models with rich controls — in particular, prior scores — yield approximately unbiased estimates of teacher effects (e.g. Chetty, Friedman and Rockoff, 2014*a*; Bacher-Hicks et al., 2019).

Given that our main outcomes of interest are mental health diagnoses and care utilization, we also assess whether students with a higher baseline propensity for mental health problems are systematically sorted to higher or lower well-being VA teachers. Specifically, we test for sorting based on parental psychiatric diagnosis and the student's own prior psychiatric diagnosis. Figure 9 shows that, after conditioning on our control vector, neither predictor is meaningfully related to well-being VA. The coefficients are close to zero in both cases, indicating that students at higher baseline mental health risk are not sorting to systematically better or worse well-being VA teachers.

[11] For comparison, Chetty, Friedman and Rockoff (2014*a*) report a coefficient of 0.022 (s.e. 0.002), and Gilraine and Pope (2021) report a coefficient of 0.028.

5.1.3 Quasi-experimental variation from staffing changes

The evidence in Section 5.1.2 does not rule out the possibility that students are sorted to teachers based on unobservable characteristics that are orthogonal to maternal income and twice-lagged outcomes. Therefore, we use quasi-experimental variation that leverages staffing changes among the full distribution of teachers to obtain an estimate of forecast bias. To do so, we construct a leave-two-year-out estimate of VA for each teacher using data from all years except $t-1$ and t. We then calculate the change in mean teacher VA for each school-grade-subject-year-cell.

Panel A of Figure 10 plots changes in cohort-level outcomes against changes in mean well-being VA. We find a strong positive relationship: increases in the quality of the teaching staff predict increases in mean student well-being across cohorts. The baseline estimate exceeds one, but this relationship is sensitive to extreme observations. When trimming the top and bottom five percent of changes, the estimated coefficient is close to one, consistent with forecast unbiasedness. This sensitivity likely reflects the smaller number of cohorts and greater noise in cohort-level means relative to other work that uses longer panels.

As a further placebo test, we repeat the staffing-change analysis using predicted outcomes based on maternal income. In this case, changes in well-being VA are uncorrelated with changes in predicted outcomes, supporting the assumption that staffing changes are not systematically related to changes in student composition (Panel B of Figure 10).

Overall, while the quasi-experimental staffing-change design is necessarily less precise in our setting, due to a shorter panel than other studies and less frequent standardized testing, the full set of bias checks yields a consistent picture. Across forecast bias tests and placebo outcomes, we find little evidence of systematic sorting that would materially bias our value-added estimates.

5.2 Bias in the Long-Term Effects

Our approach to estimating teachers' long-term impacts builds on the evidence above that conditioning on prior outcomes and a rich set of observables yields approximately unbiased estimates of teachers' causal effects on contemporaneous outcomes. Under this assumption, differences in longer-run outcomes across students assigned to different teachers—conditional on the same control vector—can be interpreted as reflecting teachers' long-term impacts. We assess the validity of this approach by exploiting quasi-experimental variation in teacher value-added arising from staffing changes across grades and schools. Following Chetty, Friedman and Rockoff (2014*b*), we regress changes in school-grade-year long-term outcomes on changes in the mean teacher value-added weighted by the number of students.

Table 3 reports these quasi-experimental estimates for well-being value-added. The signs of the estimated coefficients are generally consistent with the results presented in Section 4. Despite the loss of statistical power, we continue to find a statistically significant effect of well-being value-added on high school well-being and mental health care utilization.

6 Multidimensional Teacher Quality

Recent work emphasizes that teacher quality is multidimensional, with teachers affecting students' long-run outcomes through distinct channels (Petek and Pope, 2023). While the focus of this paper is on measuring teachers' impacts on student well-being and their longer-run mental health consequences, it is natural to ask how these effects relate to more traditional measures of teacher quality. In particular, we examine two questions: (i) whether teachers who improve students' well-being also raise test scores, and (ii) whether teachers who raise test scores have spillover effects on longer-run mental health outcomes.

As reported in Table 4, the SD of teacher effects on test scores is 0.151 for Danish teachers and 0.202 for Math teachers — somewhat larger than the corresponding well-being VA estimates, and consistent with prior estimates in the literature (e.g., Chetty, Friedman and Rockoff, 2014*a*; Rose, Schellenberg and Shem-Tov, 2022). The jackknife test score VA estimates have a standard deviation of 0.079 (Figure 1). We now ask how these two dimensions of teacher quality relate to each other and to students' long-run outcomes.

6.1 Within-Teacher Correlation Between VA Measures

Because each value-added measure is measured with error, the observed correlation between the two may differ from the correlation between teachers' true effects. Random estimation error can attenuate correlations toward zero, while estimation errors that are correlated across outcomes within the same year may bias correlations upward.

To address these concerns, we implement a split-sample approach to estimate an attenuation-corrected correlation between the two VA measures, following Denning et al. (2025). We randomly split students within each teacher-year cell into two independent samples (A and B). We then estimate each teacher's VA on outcome 0 (well-being) and outcome 1 (test scores) separately in each sample, yielding four VA estimates per teacher: μ_0^A, μ_0^B, μ_1^A, μ_1^B. We then estimate the correlation as:

$$\hat{r}_{01} = \frac{c\hat{o}rr(\mu_0^A, \mu_1^B)}{\sqrt{c\hat{o}rr(\mu_0^A, \mu_0^B) c\hat{o}rr(\mu_1^A, \mu_1^B)}} \tag{7}$$

The split-sample correlation in the numerator eliminates spurious correlation arising from shared estimation error across outcomes, while the denominator corrects for attenuation due to measurement error in each value-added measure. We bootstrap this estimate with 100 random sample splits, each yielding two cross-sample estimates, and report the average of the 200 resulting estimates.

Figure 11 plots test score and well-being value-added and shows that, although both dimensions of teacher quality are positively correlated, the correlation is relatively weak. Some teachers who perform poorly as measured by test score value-added perform well on the well-being value-added dimension, and vice-versa. For example, only 27 percent of teachers in the top quintile of well-being value-added are also in the top quintile of test-score value-added.

While Figure 11 illustrates the raw relationship between the two value-added measures, this correlation may be affected by measurement error. Applying the split-sample procedure described above yields an attenuation-corrected correlation of approximately 0.25 between teachers' well-being and test score value-added. This estimate is close to the raw correlation, indicating that the weak relationship between the two measures is not driven by attenuation from estimation error.

6.2 Long-Run Effects of Test Score Value-Added

The weak correlation between well-being and test score value-added suggests that these measures may be capturing distinct dimensions of teacher quality.[12] We therefore next examine whether teachers' test score value-added has independent long-run effects on students' mental health outcomes, and how these effects compare to those associated with well-being value-added. We proceed in two steps. First, we estimate the long-run effects of test score value-added on mental health and academic outcomes in univariate specifications. We then estimate models that include both test score and well-being value-added jointly, allowing us to assess the extent to which each dimension independently predicts students' long-term outcomes.

[12] We conduct the same forecast bias, selection-on-observables, and quasi-experimental staffing-change checks for test score value-added as for well-being value-added. Results are reported in Appendix Figures A.2 and A.3 and are broadly consistent with unbiased estimates.

Appendix Figure A.4 show that teachers with higher test score value-added modestly increase students' high school well-being, though the magnitude is small relative to the corresponding effect of well-being value-added: a one–standard deviation increase in test score VA raises high school well-being by 0.0015 standard deviations, compared to 0.044 for well-being VA. Test score value-added has no statistically significant effect on mental health care utilization but is associated with a reduction in the likelihood of receiving a psychiatric diagnosis, driven primarily by anxiety diagnoses (Appendix Figure A.5). Notably, these diagnosis effects are even larger than those associated with well-being value-added, suggesting that different dimensions of teacher quality may influence mental health through distinct channels.

Turning to academic outcomes, Appendix Figure A.6 show that test score value-added has substantially larger effects. A one standard deviation increase in test score VA raises the probability of enrolling in an academic high school track by 0.016 and increases standardized high school GPA by 0.054. These effects are naturally larger than those associated with well-being value-added, indicating that test score value-added is a strong predictor of long-run academic success.[13]

6.3 Joint Effects of Well-Being and Test Score Value-Added

We next estimate specifications that include both well-being and test score value-added simultaneously, allowing us to assess whether each dimension predicts long-run outcomes independently. Figure 12a show that, for high school well-being, the effect of well-being value-added remains large and statistically significant: a one–standard deviation increase in well-being VA raises high school well-being by 0.048 standard deviations, while the coefficient on test score value-added becomes small and statistically insignificant.

Figure 13 extends this analysis to the five high school well-being sub-indices. Well-being value-added improves all five dimensions uniformly. Test score value-added, by contrast, has a positive and significant effect only on academic well-being, which captures students' engagement, motivation, self-efficacy, and perceived academic performance. This suggests that teachers who raise test scores in middle school not only improve academic achievement in high school, but also improve students' academic well-being, reflecting a coherent academic dimension of teacher quality that shows up in both performance and in how students

[13]We validate these estimates using quasi-experimental variation from staffing changes across cohorts. The quasi-experimental estimates for test score value-added are broadly consistent with the reduced-form results, with similar directional patterns across outcomes; see Appendix Table A.1.

experience school. At the same time, test score value-added has null or even slightly negative effects on the remaining dimensions — social well-being, learning environment, pressure and concerns, and bullying — suggesting that the gains associated with high test score VA teachers may come partly at the expense of students' social and emotional experience of school. This provides an illustration of why the two dimensions capture meaningfully distinct aspects of teacher quality.

For mental health care utilization, Figure 12b shows that well-being value-added continues to significantly reduce the probability of a visit, whereas test score value-added is associated with a small increase in utilization. Turning to psychiatric diagnoses, Figure 12c shows that both dimensions remain relevant: test score value-added is associated with a significant reduction in diagnoses, driven primarily by anxiety, while the aggregate effect of well-being value-added is smaller and imprecisely estimated, though diagnosis-specific patterns persist—well-being value-added continues to predict reductions in depression, while test score value-added predicts reductions in anxiety (Figure 14). This pattern is consistent with a view in which social and relational dimensions of school life are protective against depression, while academic pressure is more closely linked to anxiety — precisely the dimensions on which the two VA types diverge in Figure 13. Results for academic outcomes in the joint specifications are reported in Appendix Figure A.7.

Taken together, these results further highlight the multidimensional nature of teacher quality. Test score and well-being value-added are only weakly correlated and predict different long-run outcomes. Well-being value-added is the primary predictor of students' long-run well-being and mental health care utilization, while test score value-added strongly predicts academic outcomes and is associated with reductions in specific psychiatric diagnoses, particularly anxiety. These patterns persist when both dimensions are included jointly, indicating that neither measure simply proxies for the other and that evaluations based solely on test score value-added miss important dimensions of teachers' impacts on students' long-run outcomes.

7 Conclusions

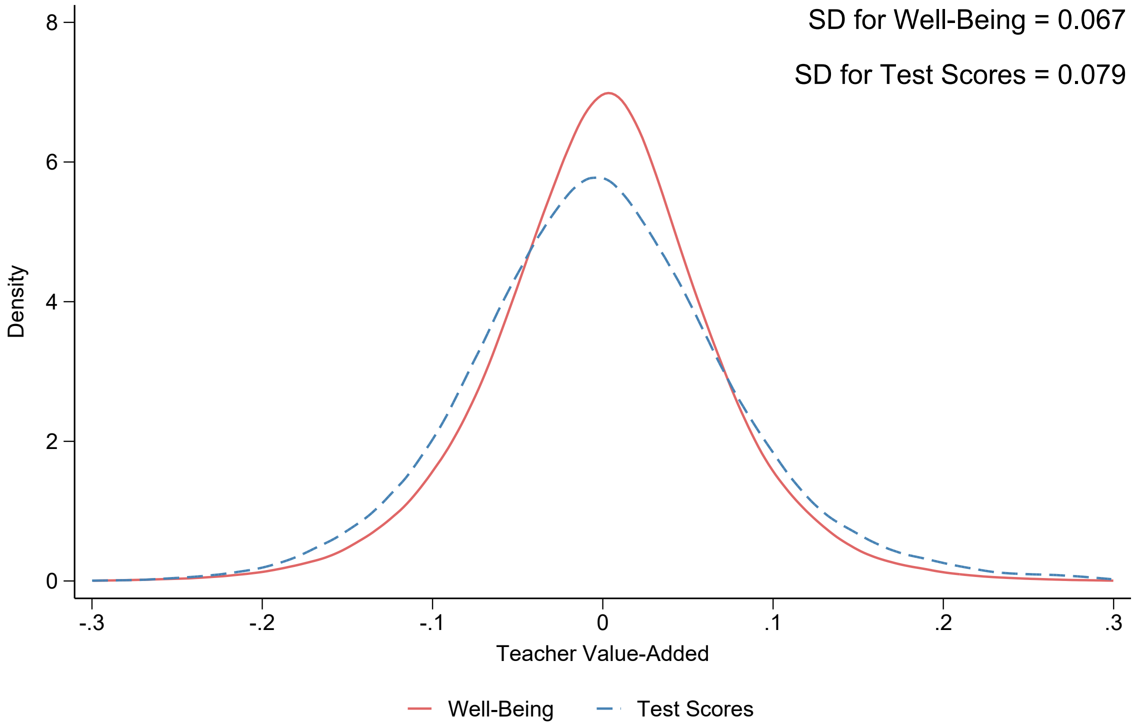
This paper shows that teachers play a meaningful role in shaping not only students' academic achievement but also their well-being and mental health. Using linked administrative and survey data from Denmark, we document substantial heterogeneity in teachers' impacts on student well-being during middle school. These effects are comparable in magnitude to teacher effects on other non–test score outcomes and are reflected on other outcomes several years later, influencing students' well-being in high school, mental health care utilization, and psychiatric diagnoses.

Our results also highlight the multidimensional nature of teacher quality. Teachers who are effective at improving student well-being are often not the same teachers who raise test scores, and these dimensions predict different long-run outcomes. Well-being value-added is the primary predictor of students' subsequent well-being and mental health care utilization, while test score value-added strongly predicts academic outcomes and reductions in specific diagnoses such as anxiety. Evaluations of teacher effectiveness that rely solely on test scores therefore miss important channels through which teachers shape students' long-run trajectories.

More broadly, our findings suggest that schools and teachers in particular are a key environment for influencing child and adolescent mental health, a growing concern with key implications for adult outcomes in the labor market. Our results suggest that incorporating student well-being into the analysis of teacher effectiveness provides a richer understanding of how education policies and teachers affect human development.

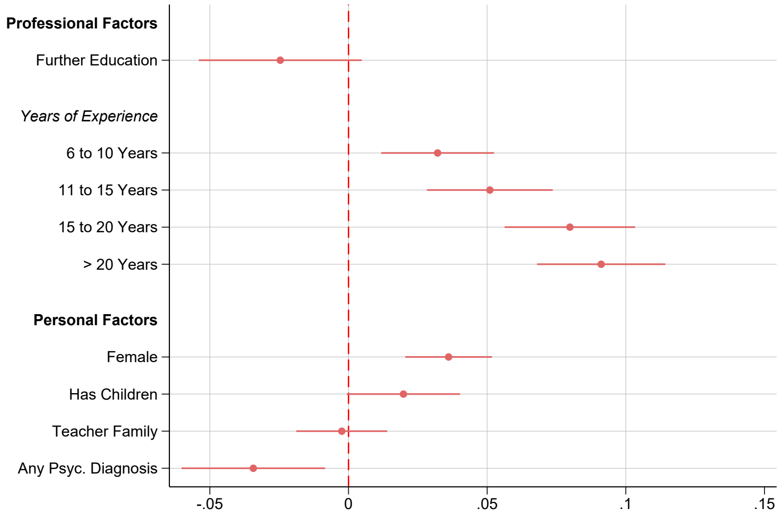
Figures

Figure 1: Empirical Distributions of Teacher VA Estimates



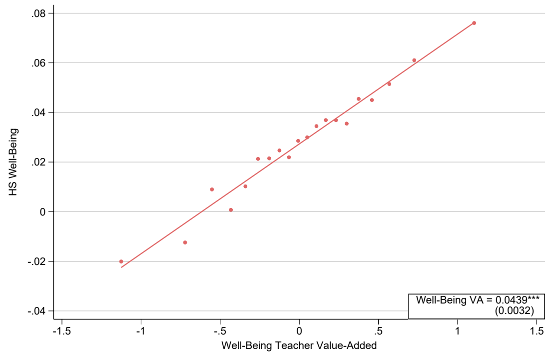
Notes: This figure plots kernel densities of the empirical distribution of teacher VA estimates on well-being and test scores, pooled across subjects. We also report the standard deviations of these empirical distributions of VA estimates. Note that these standard deviations are smaller than the standard deviation of true teacher effects reported in Table 4 because VA estimates are shrunk toward the mean to account for noise and obtain unbiased forecasts.

Figure 2: Correlates of Well-Being Value-Added

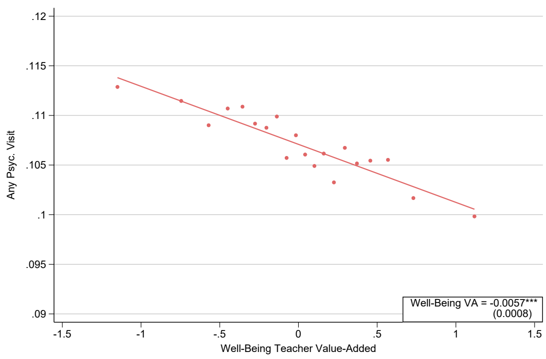


Notes: This figure reports estimates of the association between teacher well-being value-added and a set of observable professional and personal teacher characteristics. Each coefficient is from a regression of the teacher's well-being VA estimate on the indicated characteristic, estimated at the teacher-year level with standard errors clustered at the teacher level. Professional factors include an indicator for further education beyond a teaching degree and indicators for years of experience relative to fewer than six years (the omitted category). Personal factors include indicators for being female, having children, coming from a teacher family, and having any psychiatric diagnosis.

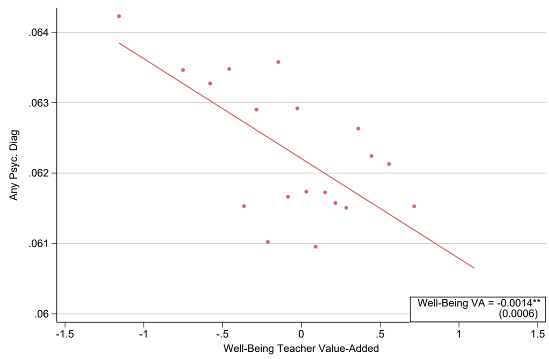
Figure 3: Effects of Teacher Well-Being Value-Added on Well-Being and Mental Health



(a) High School Well-Being



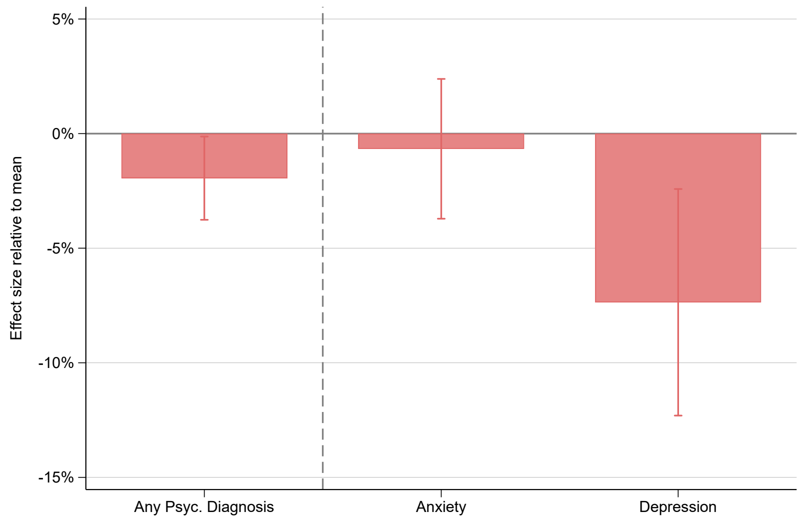
(b) Mental Health Care Utilization (Ages 16–18)



(c) Psychiatric Diagnosis (Ages 16–18)

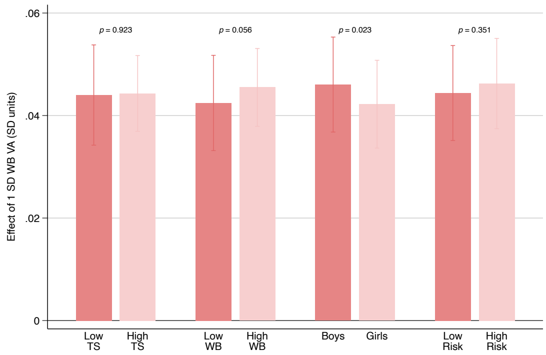
Notes: This figure plots the relationship between teacher well-being value-added in middle school and students' well-being and mental health outcomes. In particular, we focus on three outcomes: (i) self-reported well-being during the first year of high school, (ii) indicator that takes value one if the student visited a psychiatrist or psychologist between ages 16 to 18 and (ii) indicator that takes value one if the student was diagnosed with any psychiatric condition between ages 16 to 18. The unit of observation is a student–subject–year. Both the outcome and the teacher value-added measure are residualized with respect to the baseline control vector used in the value-added estimation. Teacher value-added estimates are divided into 20 equal-sized bins, and the figure plots the mean residualized outcome within each bin against the mean value-added estimate, with the unconditional means added back. We also plot the best linear fit estimated before binning the data and report its slope coefficient and standard error, clustered at the student and classroom level.

Figure 4: Effects of Well-Being Value-Added on Psychiatric Diagnoses

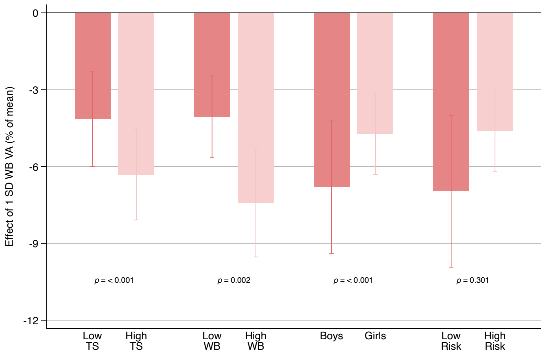


Notes: This figure reports estimates of the effects of well-being value-added on the probability of receiving specific psychiatric diagnoses between ages 16 and 18. The value-added measure is standardized and constructed using leave-year-out estimates. All regressions include the baseline control vector used in the value-added estimation. Standard errors are clustered at the student and classroom level.

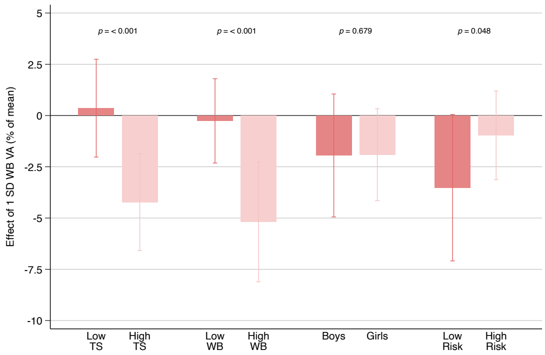
Figure 5: Heterogeneous Effects



(a) High School Well-Being



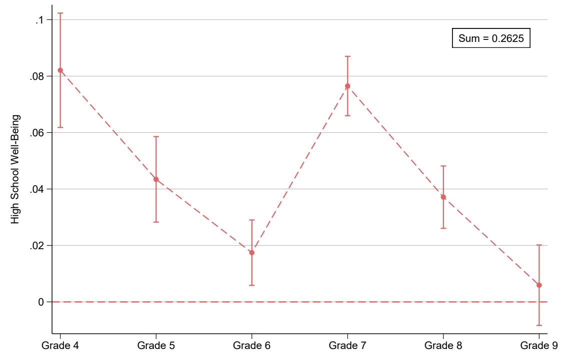
(b) Mental Health Care Utilization (Ages 16–18)



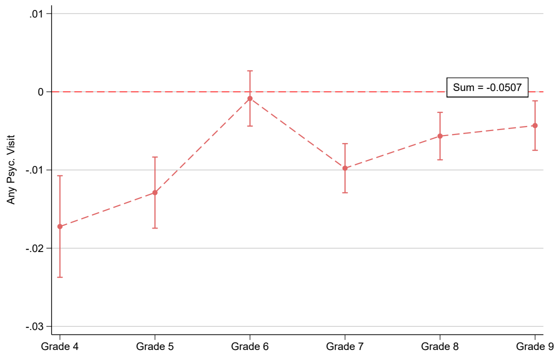
(c) Psychiatric Diagnosis (Ages 16–18)

Notes: These figures report heterogeneous effects of WB teacher value-added on high school well-being, mental health care utilization, and psychiatric diagnoses by student gender, prior academic achievement, prior well-being, and predicted risk of mental health care utilization. Each panel presents coefficients from split-sample regressions estimated separately for the indicated subgroup. Teacher value-added is standardized, and all specifications include the baseline control vector used in the value-added estimation. Effects in Panel A are in standard deviation units; effects in Panels B and C are expressed as a percentage of the subgroup mean. The p-values reported correspond to a test of equality of coefficients across subgroups, estimated from a pooled regression interacting teacher value-added with the subgroup indicator.

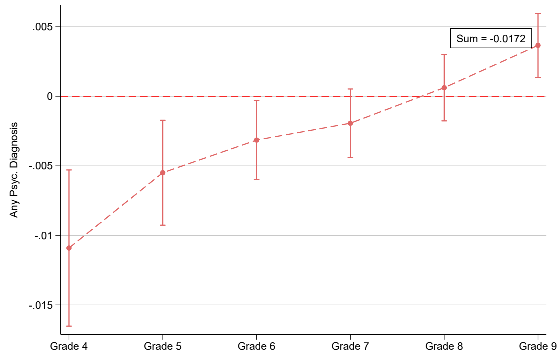
Figure 6: Effects of Well-Being Value-Added by Grade



(a) High School Well-Being



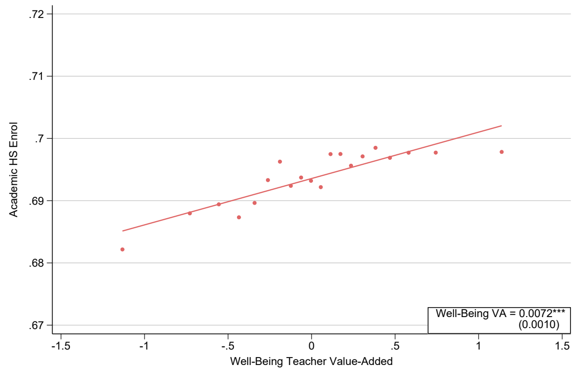
(b) Mental Health Care Utilization (Ages 16–18)



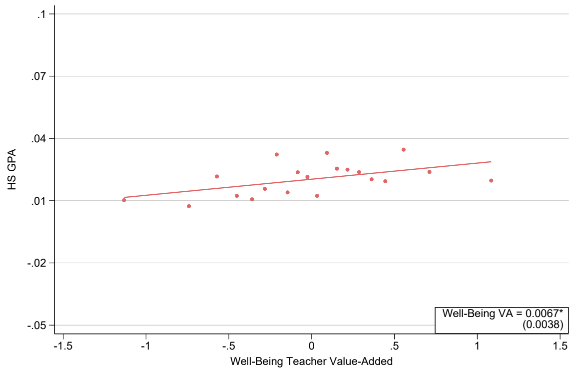
(c) Psychiatric Diagnosis (Ages 16–18)

Notes: This figure plots the effect of well-being value-added on three outcomes by the grade level of the student: self-reported well-being during the first year of high school (Panel A), an indicator for visiting a psychiatrist or psychologist between ages 16 and 18 (Panel B), and an indicator for receiving a psychiatric diagnosis between ages 16 and 18 (Panel C). The plotted coefficients and 95 percent confidence intervals are from regressions of each outcome on well-being value-added and the baseline control vector used in the value-added estimation, estimated separately for each grade. Standard errors are clustered at the classroom level. Each panel also reports the sum of coefficients across all grades, which under the assumption of no diminishing returns provides an upper bound on the cumulative effect of having a one standard deviation higher teacher VA in every grade from 4 to 9.

Figure 7: Effects of Teacher Well-Being Value-Added on Academic Outcomes



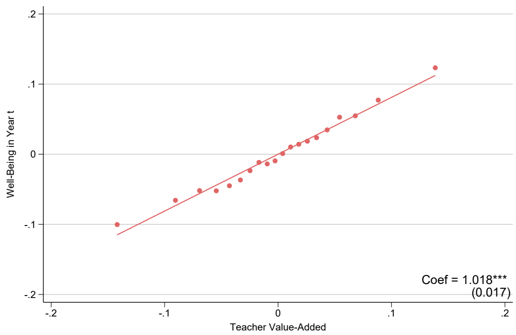
(a) Academic HS Enrollment



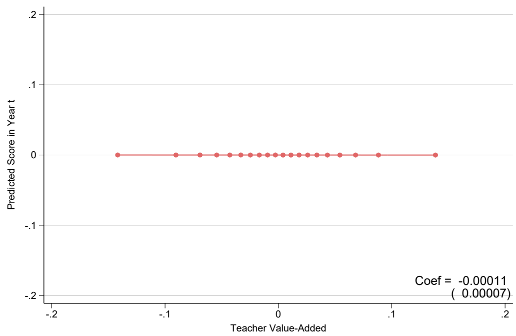
(b) High School GPA

Notes: This figure plots the relationship between teachers' well-being value-added in middle school and students' academic outcomes. In particular, we focus on two outcomes: (i) indicator that takes value one if the student enrolled in academic high school and (ii) high school GPA, standardized by graduation year and academic track. The unit of observation is a student–subject–year. Both the outcome and the teacher value-added measure are residualized with respect to the baseline control vector used in the value-added estimation. Teacher value-added estimates are divided into 20 equal-sized bins, and the figure plots the mean residualized outcome within each bin against the mean value-added estimate, with the unconditional means added back. We also plot the best linear fit estimated before binning the data and report its slope coefficient and standard error, clustered at the student and classroom level.

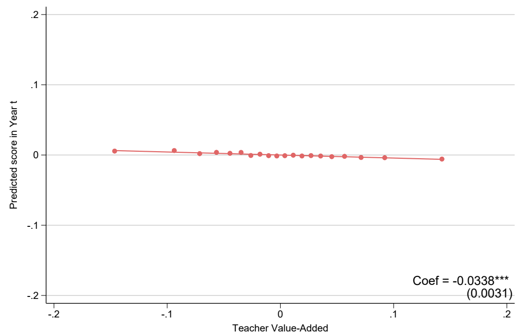
Figure 8: Effects of Well-Being Teacher Value-Added on Actual and Predicted Scores



(a) Actual Score



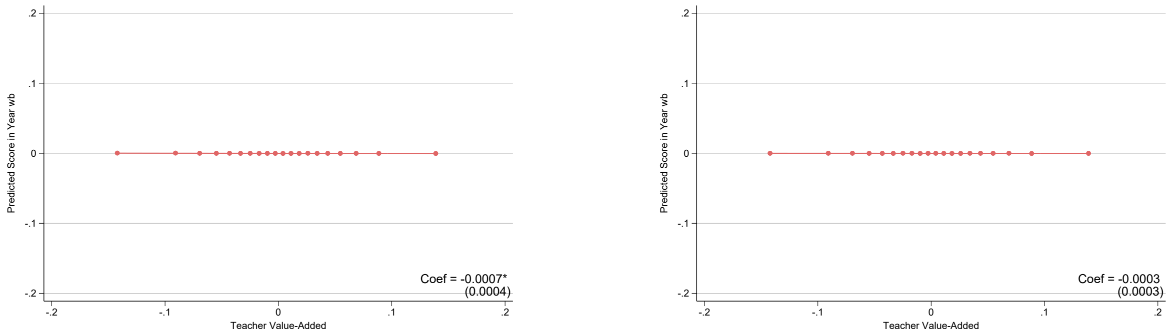
(b) Predicted Score Using Maternal Income



(c) Predicted Score Using Twice-Lagged Well-Being

Notes: This figure assesses the validity of the well-being value-added estimates. Panel A plots contemporaneous residualized well-being against teacher well-being value-added. Under the stationarity assumption, the coefficient should equal one. Panels B and C assess whether students sort to teachers based on observable characteristics excluded from the value-added model. Panel B plots predicted outcomes based on maternal income; Panel C plots predicted outcomes based on twice-lagged well-being. Predicted outcomes are obtained by residualizing the predictor with respect to the baseline control vector and regressing residualized outcomes on this residualized predictor. In all panels, teacher value-added estimates are divided into 20 equal-sized bins and the figures plot the mean outcome within each bin against the mean value-added estimate. Solid lines show best linear fits estimated on the underlying microdata using OLS. Standard errors are clustered at the school-cohort level.

Figure 9: Teacher Well-Being Value-Added and Predicted Outcomes Based on Psychiatric Diagnoses

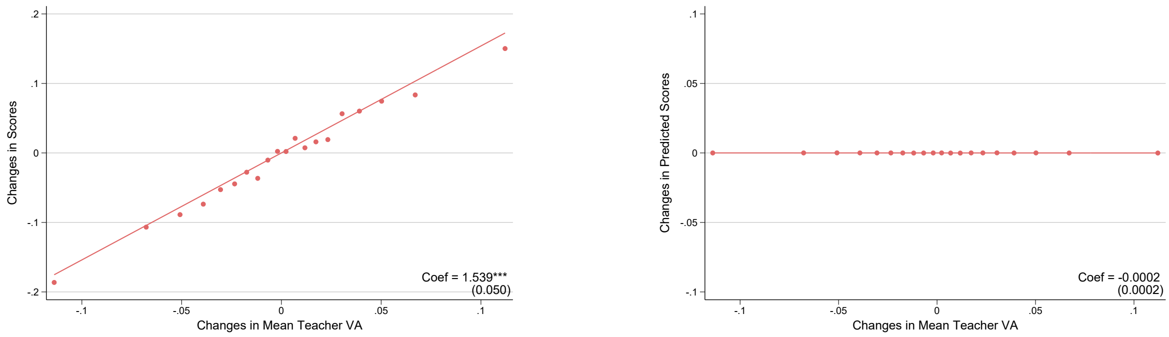


(a) Parental Psychiatric Diagnosis

(b) Prior Student Psychiatric Diagnosis

Notes: This figure assesses whether students sort to well-being VA teachers based on baseline mental health risk. Panel A plots predicted outcomes based on parental psychiatric diagnosis, and Panel B plots predicted outcomes based on the student's own prior psychiatric diagnosis. Predicted outcomes are obtained by residualizing each predictor with respect to the baseline control vector and regressing residualized well-being on these residualized predictors. Teacher value-added estimates are divided into 20 equal-sized bins, and the figures plot the mean predicted outcome within each bin against the mean value-added estimate. Solid lines show best linear fits estimated on the underlying microdata using OLS. Standard errors are clustered at the school-cohort level.

Figure 10: Effects of Changes in Teaching Staff on Student Outcomes Across Cohorts: Well-Being Value-Added

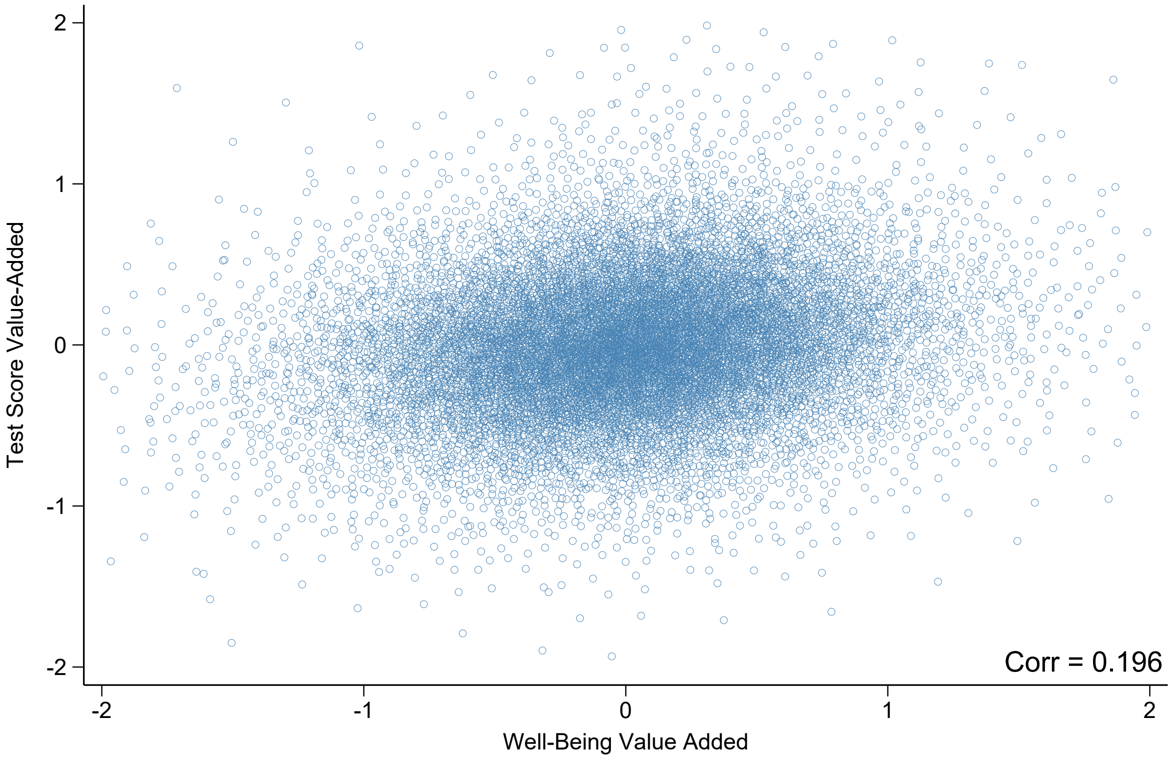


(a) Actual Outcomes

(b) Predicted Outcomes (Placebo)

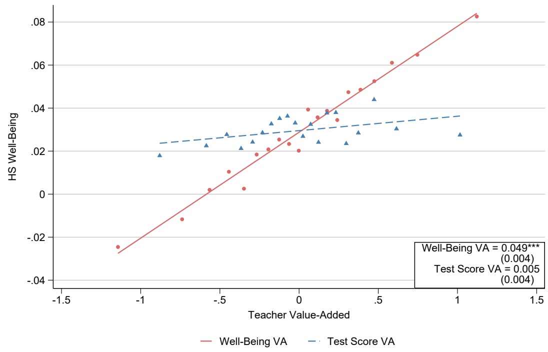
Notes: This figure plots changes in cohort-level mean outcomes against changes in mean well-being value-added within school–grade–subject cells, exploiting staffing changes across cohorts. Panel A uses actual outcomes, while Panel B uses predicted outcomes based on maternal income as a placebo check. Teacher value-added is constructed using leave-two-year-out estimates. Lines show best linear fits estimated on the underlying microdata.

Figure 11: Within-Teacher Correlation Between VA Measures

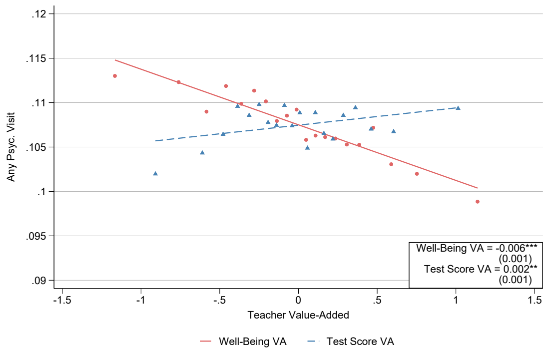


Notes: This scatter plot shows a plot of teachers' normalized test score and well-being value-added within 2 standard deviations of the mean. For confidentiality reasons, Each point shows the mean of 6 nearest-neighbor observations.

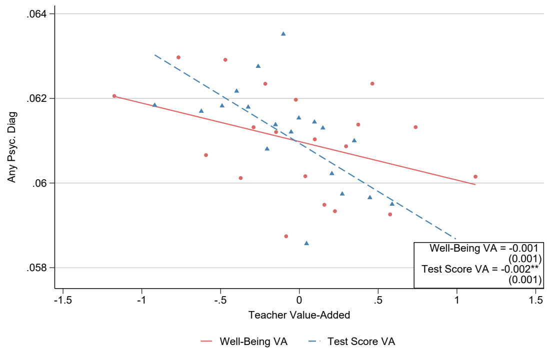
Figure 12: Effects of Teacher Value-Added on Well-Being and Mental Health (Multivariate)



(a) High School Well-Being



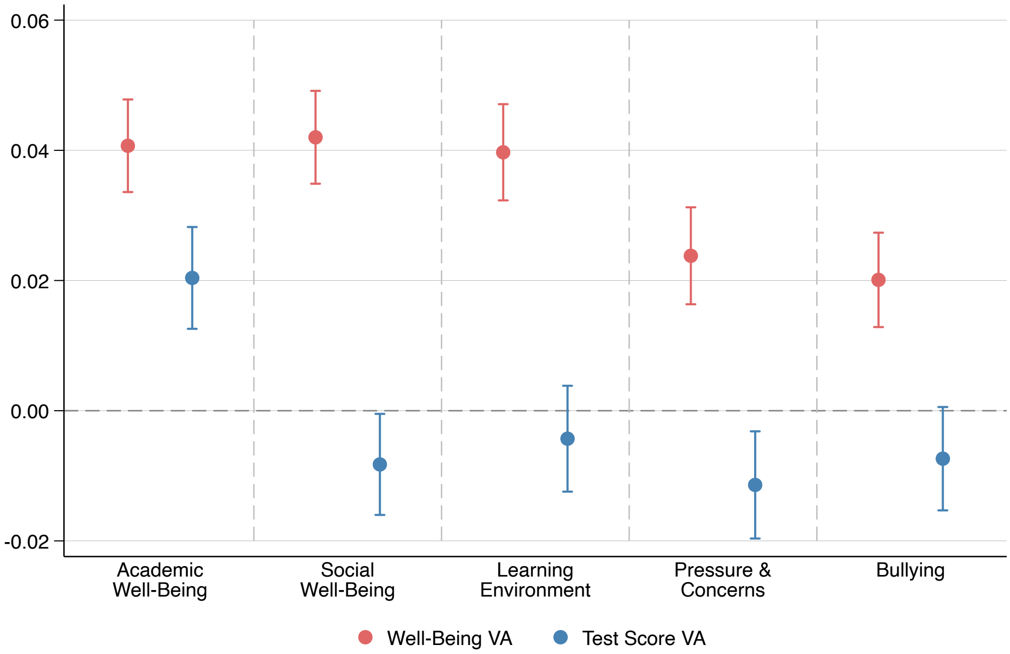
(b) Mental Health Care Utilization (Ages 16–18)



(c) Psychiatric Diagnosis (Ages 16–18)

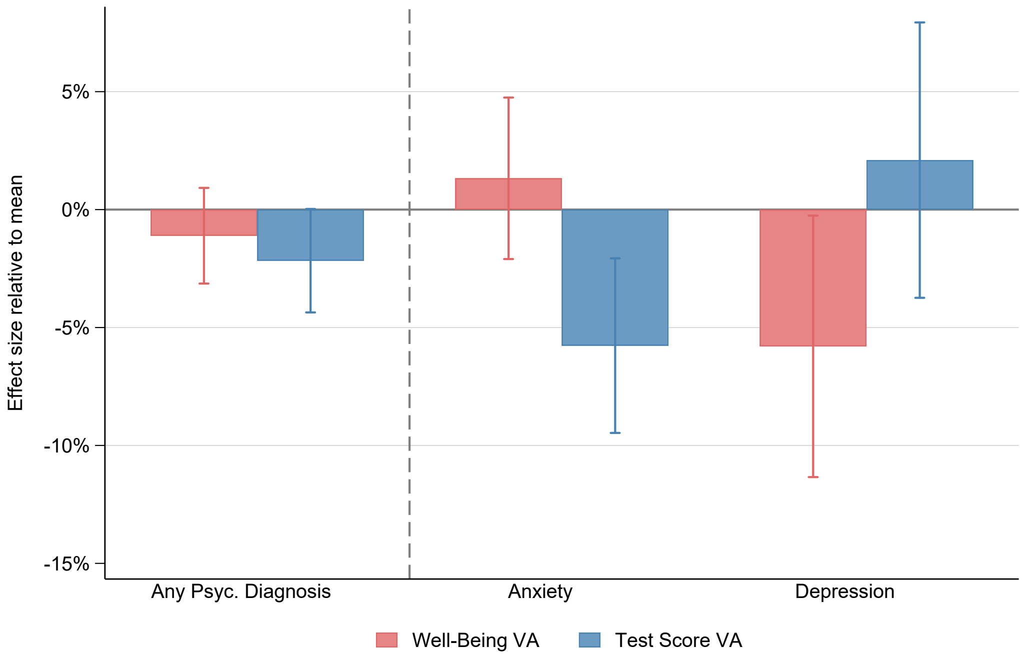
Notes: This figure plots the partial relationship between teachers' value-added in middle school and students' well-being and mental health outcomes. In particular, we focus on three outcomes: (i) self-reported well-being during the first year of high school, (ii) indicator that takes value one if the student visited a psychiatrist or psychologist between ages 16 to 18 and (ii) indicator that takes value one if the student was diagnosed with any psychiatric condition between ages 16 to 18. The figure displays two series corresponding to teacher well-being value-added and test score value-added, respectively; both series are based on a single multivariate specification in which the outcome is regressed simultaneously on both value-added measures. The unit of observation is a student–subject–year. Both the outcome and the teacher value-added measure are residualized with respect to the baseline control vector used in the value-added estimation. Teacher value-added estimates are divided into 20 equal-sized bins, and the figure plots the mean residualized outcome within each bin against the mean value-added estimate, with the unconditional means added back. We also plot the best linear fit estimated before binning the data and report its slope coefficient and standard error, clustered at the student and classroom level.

Figure 13: Effects of Teacher Value-Added on High School Well-Being Sub-Indices



Notes: This figure reports estimates of the effects of well-being value-added and test score value-added on the five high school well-being sub-indices constructed by the Ministry of Education: academic well-being, social well-being, learning environment, pressure and concerns, and bullying. The headline high school well-being measure used in the main analysis is the average of the academic and social well-being sub-indices. Each coefficient is from a separate regression of the standardized sub-index on the indicated value-added measure, conditional on the baseline control vector used in the value-added estimation. Error bars indicate 95 percent confidence intervals, with standard errors clustered at the student and classroom level.

Figure 14: Joint Effects of Teacher Value-Added on Psychiatric Diagnoses



Notes: This figure reports estimates of the joint effects of well-being and test score value-added on the probability of receiving specific psychiatric diagnoses between ages 16 and 18. Both value-added measures are standardized and constructed using leave-year-out estimates. All regressions include the baseline control vector used in the value-added estimation. Standard errors are clustered at the student and classroom level.

Tables

Table 1: Descriptive Statistics

	Full Sample			Restricted Sample		
	Mean	SD	Obs.	Mean	SD	Obs.
A. Demographics						
Female	0.49	0.50	1,921,439	0.50	0.50	1,007,038
Age	12.14	1.79	1,921,439	12.00	1.76	1,007,038
Migration Background	0.12	0.32	1,921,439	0.11	0.32	1,007,038
Mother's Education	0.40	0.49	1,872,611	0.42	0.49	986,414
Father's Education	0.27	0.45	1,816,746	0.29	0.45	960,107
B. Short-Run Outcomes						
Well-Being Score	0.02	0.99	1,421,968	0.04	0.98	1,007,100
Danish Test Score	0.05	0.95	1,090,303	0.09	0.92	542,643
Math Test Score	0.04	0.97	635,783	0.10	0.94	330,420
C. Long-Run Outcomes						
High School Enrollment	0.64	0.48	1,923,690	0.61	0.49	1,007,100
High School GPA	0.00	1.00	738,828	0.02	1.00	272,653
High School Well-Being	0.00	0.99	757,174	0.03	0.98	425,630
Any Criminal Charge	0.09	0.29	1,644,555	0.08	0.26	792,051
Any Visit Psychologist	0.11	0.31	1,641,728	0.10	0.30	789,984
Any Visit Psychiatrist	0.02	0.15	1,639,018	0.02	0.14	787,877
Any Psychiatric Diagnosis	0.07	0.26	1,454,829	0.06	0.24	646,031

Notes: This table presents summary statistics for demographics, short-run outcomes, and long-run outcomes in the full sample and the restricted analysis sample, as described in Section 2. Parental education variables take value 1 if the parent has a tertiary degree. Well-being and test scores are standardized by cohort. Long-run outcomes unrelated to high school are measured between ages 16 and 18. In each analysis, we use the largest sample possible given when an outcome is observed.

Table 2: Relationship between Short-Run Measures and Long-Term Outcomes

	HS Well-Being	Psyc. Visit	Psyc. Diag.	HS Enrol	HS GPA
Well-Being	0.289***	-0.0271***	-0.0253***	0.0341***	0.0589***
	(88.40)	(-36.12)	(-36.67)	(40.70)	(19.72)
Test Scores	0.0194***	-0.00161**	-0.00422***	0.0888***	0.354***
	(6.93)	(-2.41)	(-6.93)	(83.36)	(111.54)
Observations	366662	660722	571054	733809	282577
Mean	0.028	0.110	0.073	0.739	0.024
R-squared	0.192	0.046	0.036	0.289	0.349

Notes: This table reports regressions of long-run student outcomes on contemporaneous measures of student performance and well-being measured during middle school. Outcomes include high school well-being, mental health care utilization, psychiatric diagnoses, and academic outcomes. All specifications include the same baseline control vector used in the value-added analysis, including lagged outcomes, student demographics, classroom and school controls, and grade-by-year fixed effects. Standard errors are clustered at the student and classroom level.

Table 3: Quasi-Experimental Estimates: Effects of Well-Being Value-Added on Long-Run Outcomes

	HS Well-Being	Psyc. Visit	Psyc. Diag.	HS Enrol	HS GPA
Well-Being VA	0.0950***	-0.00869***	-0.00639***	0.0122***	0.0288**
	(9.83)	(-2.98)	(-2.74)	(4.09)	(2.22)
Observations	22,396	22,645	22,878	28,967	14,008
R-squared	0.007	0.001	0.001	0.063	0.001

Notes: This table exploits quasi-experimental variation in teacher quality generated by staffing changes across adjacent cohorts within school–grade cells. The unit of observation is a school–grade–cohort. Each column reports the effect of a one standard deviation increase in well-being value-added on cohort-level high school outcomes, using within school–grade variation in teacher quality. Specifically, we regress changes in the school–grade–cohort mean of each outcome on changes in the corresponding school–grade–cohort mean of teacher value-added, estimated using leave-two-year-out measures of teacher quality. All specifications include year fixed effects and are weighted by cohort size. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 4: Variance Estimates

	Well-Being			Test Scores		
	Danish	Math	Pooled	Danish	Math	Pooled
Total SD	–	–	0.774	0.628	0.614	0.627
Individual-level SD	–	–	0.741	0.585	0.564	0.582
Class-level SD	–	–	0.194	0.173	0.135	0.150
Teacher SD	0.120	0.101	0.109	0.151	0.202	0.181
Estimate Teacher SD	0.134	0.094	0.108	0.129	0.144	0.132

Notes: This table reports the standard deviation of outcome residuals and decomposes the total variation into individual-, class-, and teacher-level components for test scores and well-being outcomes in Danish and Mathematics. The first row reports the total standard deviation of residuals. Rows 2–4 decompose this variation into components attributable to individual-level variation, class-level shocks, and teacher-level variation, respectively. The variances corresponding to the individual-, class-, and teacher-level components sum to the total variance. In Row 4, we estimate the standard deviation of teacher effects as the square root of the covariance of mean score residuals across a random pair of classrooms within the same year. In cases where teachers are observed teaching only one classroom per year and class-level and teacher-level shocks cannot be separately identified, we report a lower bound on the within-year standard deviation of teacher effects, computed as the square root of the autocovariance of classroom residuals across a one-year lag (Row 5).

References

Adhvaryu, Achyuta, James Fenske, and Anant Nyshadham. 2019. "Early life circumstance and adult mental health." *Journal of Political Economy*, 127(4): 1516–1549.

Bacher-Hicks, Andrew, Mark J. Chin, Thomas J. Kane, and Douglas O. Staiger. 2019. "An experimental evaluation of three teacher quality measures: Value-added, classroom observations, and student surveys." *Economics of Education Review*, 73: 101919.

Backes, Ben, James Cowan, Dan Goldhaber, and Roddy Theobald. 2024. "How to measure a teacher: The influence of test and nontest value-added on long-run student outcomes." *Journal of Human Resources*.

Baird, Sarah, Jacobus De Hoop, and Berk Özler. 2013. "Income shocks and adolescent mental health." *Journal of Human Resources*, 48(2): 370–403.

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

Beuchert, Louise V, and Anne B Nandrup. 2018. "The Danish national tests at a glance." *Nationaløkonomisk tidsskrift*, 2: 1–37.

Bütikofer, Aline, Rita Ginja, Krzysztof Karbownik, and Fanny Landaud. 2023. "(Breaking) intergenerational transmission of mental health." *Journal of Human Resources*. Publisher: University of Wisconsin Press Section: Articles.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014*a*. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review*, 104(9): 2593–2632.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014*b*. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review*, 104(9): 2633–2679.

Currie, Janet. 2024. "The economics of child mental health: Introducing the causes and consequences of child mental health special issue." *Journal of Human Resources*, 59(S): S1–S13.

Dalsgaard, Søren, Erla Thorsteinsson, Betina B Trabjerg, Jörg Schullehner, Oleguer Plana-Ripoll, Isabell Brikell, Theresa Wimberley, Malene Thygesen, Kathrine Bang Madsen,

Allan Timmerman, et al. 2020. "Incidence rates and cumulative incidences of the full spectrum of diagnosed mental disorders in childhood and adolescence." *JAMA psychiatry*, 77(2): 155–164.

Danish Ministry of Education. 2016. "Methods Brief: Calculating the Well-being Indicators (Metodenotat om beregning af trivselsindikatorer, 2014/2015)." *https://www.uvm.dk/-/media/filer/uvm/stat/pdf16/160316-metodenotat-vedr-indikatorberegningerne-24022016.pdf*, Accessed: 2025-10-03.

Danish Ministry of Education. 2023. "Methods Brief: Factor Analysis and Calculating the Well-being Indicators (Metodebeskrivelse – faktoranalyse og indikatorfastlæggelse for den gymnasiale trivselsmåling)." *hhttps://www.uvm.dk/-/media/filer/uvm/udd/gym/pdf20/jan/200123-metodenotat-indikatorberegning-gymnasial-trivsel.pdf, note = Accessed: 2025-10-03*.

Dee, Thomas S, and Hans Henrik Sievertsen. 2018. "The gift of time? School starting age and mental health." *Health economics*, 27(5): 781–802.

Denning, Jeff, Rachel Nesbit, Nolan Pope, and Merrill Warnick. 2025. "Easy A's, Less Pay: The Long-Term Effects of Grade Inflation."

Fleche, Sarah. 2017. "Teacher quality, test scores and non-cognitive skills: Evidence from primary school teachers in the UK."

Gilraine, Michael, and Nolan G Pope. 2021. "Making teaching last: Long-run value-added." National Bureau of Economic Research.

Hansen, Benjamin, Joseph J Sabia, and Jessamyn Schaller. 2024. "In-person schooling and youth suicide: Evidence from school calendars and pandemic school closures." *Journal of Human Resources*, 59(S): S227–S255.

Jackson, C. Kirabo. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes." *Journal of Political Economy*, 126(5): 2072–2107. Publisher: The University of Chicago Press.

Jackson, C. Kirabo, Shanette C. Porter, John Q. Easton, Alyssa Blanchard, and Sebastián Kiguel. 2020. "School Effects on Socioemotional Development, School-Based Arrests, and Educational Attainment." *American Economic Review: Insights*, 2(4): 491–508.

Koedel, Cory, Kata Mihaly, and Jonah E. Rockoff. 2015. "Value-added modeling: A review." *Economics of Education Review*, 47: 180–195.

Kraft, Matthew A. 2019. "Teacher Effects on Complex Cognitive Skills and Social-Emotional Competencies." *Journal of Human Resources*, 54(1): 1–36.

Marcus, Jan, Simon Reif, Amelie Wuppermann, and Amélie Rouche. 2020. "Increased instruction time and stress-related health problems among school children." *Journal of Health Economics*, 70: 102256.

Montonen, Jerry. 2025. "Shaping Young Minds: The Effects of Teachers on Socio-Emotional Skills." Unpublished manuscript.

Persson, Petra, and Maya Rossin-Slater. 2018. "Family ruptures, stress, and the mental health of the next generation." *American economic review*, 108(4-5): 1214–1252.

Petek, Nathan, and Nolan G. Pope. 2023. "The Multidimensional Impact of Teachers on Students." *Journal of Political Economy*, 131(4): 1057–1107.

Rockoff, Jonah E. 2004. "The impact of individual teachers on student achievement: Evidence from panel data." *American economic review*, 94(2): 247–252.

Rose, Evan K., Jonathan T. Schellenberg, and Yotam Shem-Tov. 2022. "The Effects of Teacher Quality on Adult Criminal Justice Contact."

Rossin-Slater, Maya, Molly Schnell, Hannes Schwandt, Sam Trejo, and Lindsey Uniat. 2020. "Local exposure to school shootings and youth antidepressant use." *Proceedings of the National Academy of Sciences*, 117(38): 23484–23489.

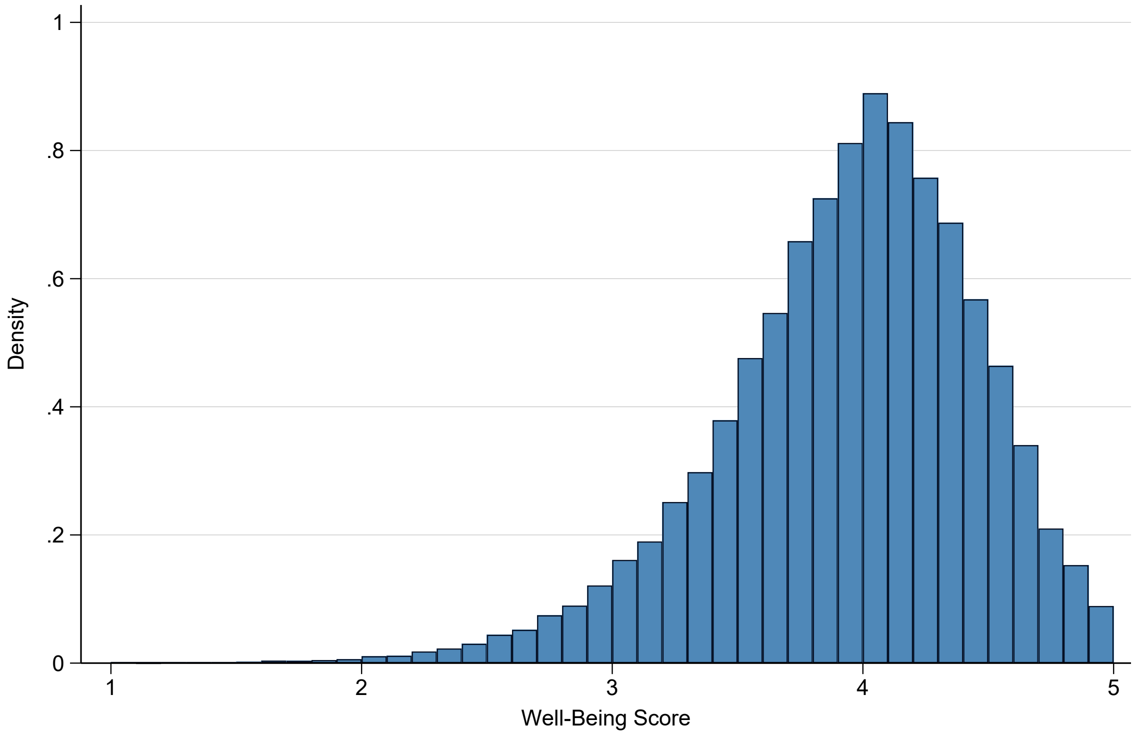
Schaller, Jessamyn, and Mariana Zerpa. 2019. "Short-run effects of parental job loss on child health." *American Journal of Health Economics*, 5(1): 8–41.

Statistics Denmark. 2024. "FORLOB10: From Primary and Lower Secondary School to Further Education." *Statistics Denmark StatBank*.

WHO. 2025. "Mental Health of Adolescents." *World Health Organization Fact Sheet*.

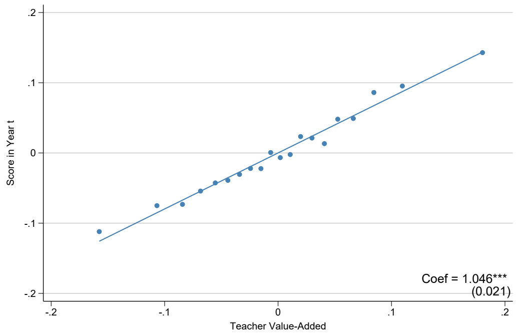
Appendix A: Additional Figures and Tables

Figure A.1: Distribution of Raw Well-Being Score

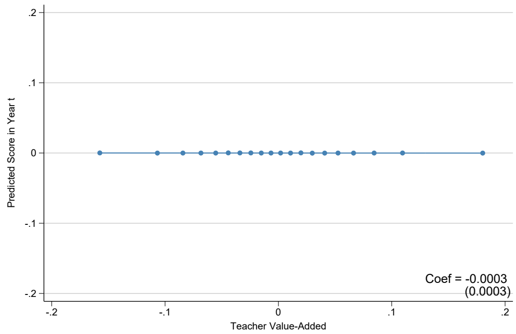


Notes: This figure shows the distribution of the raw well-being score calculated for each student in a given year. To construct the raw well-being score, we follow the definitions by the Danish Ministry of Education (2016) of academic and social well-being and compute the within-individual average of the answers provided.

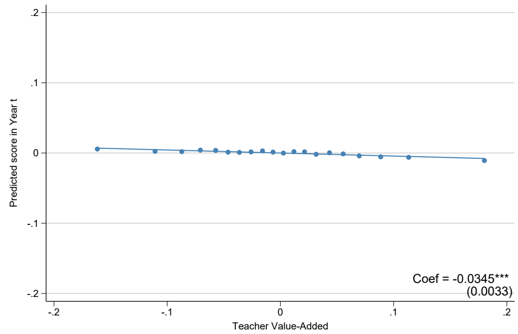
Figure A.2: Effects of Test Score Teacher Value-Added on Actual and Predicted Scores



(a) Actual Score



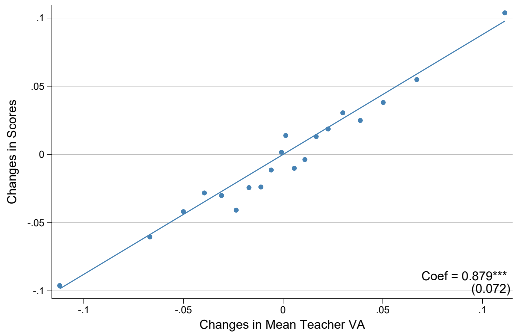
(b) Predicted Score Using Maternal Income



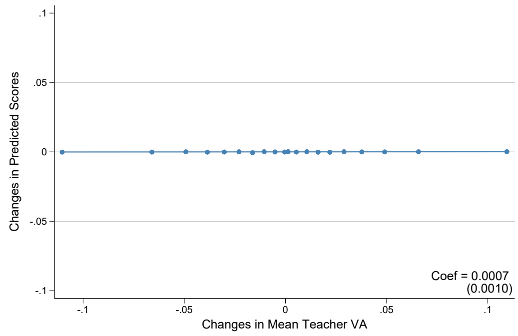
(c) Predicted Score Using Twice-Lagged Test Score

Notes: This figure assesses the validity of the test score value-added estimates. Panel A plots contemporaneous residualized test scores against teacher test score value-added. Under the stationarity assumption, the coefficient should equal one. Panels B and C assess whether students sort to teachers based on observable characteristics excluded from the value-added model. Panel B plots predicted outcomes based on maternal income; Panel C plots predicted outcomes based on twice-lagged test scores. Predicted outcomes are obtained by residualizing the predictor with respect to the baseline control vector and regressing residualized outcomes on this residualized predictor. In all panels, teacher value-added estimates are divided into 20 equal-sized bins and the figures plot the mean outcome within each bin against the mean value-added estimate. Solid lines show best linear fits estimated on the underlying microdata using OLS. Standard errors are clustered at the school-cohort level.

Figure A.3: Effects of Changes in Teaching Staff on Student Outcomes Across Cohorts: Test Score Value-Added



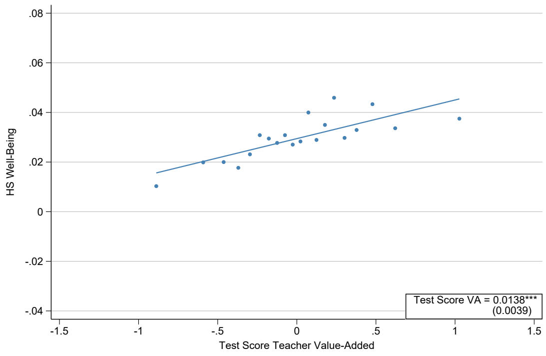
(a) Actual Outcomes



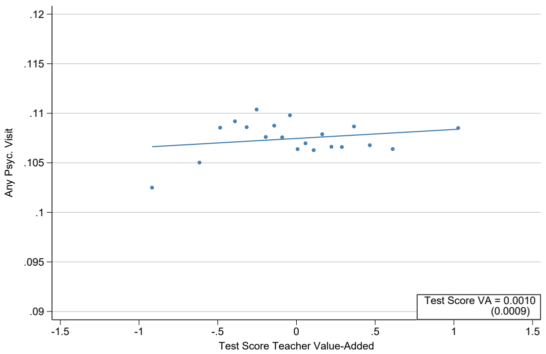
(b) Predicted Outcomes (Placebo)

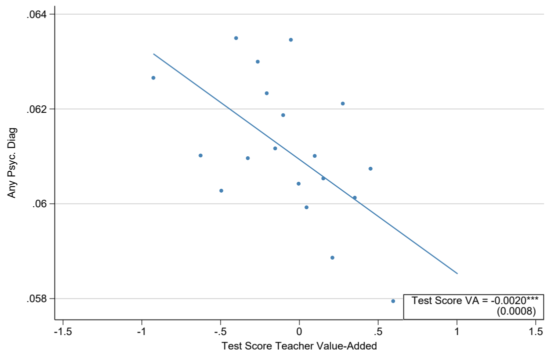
Notes: This figure plots changes in cohort-level mean outcomes against changes in mean test score value-added within school–grade–subject cells, exploiting staffing changes across cohorts. Panel A uses actual outcomes, while Panel B uses predicted outcomes based on maternal income as a placebo check. Teacher value-added is constructed using leave-two-year-out estimates. Lines show best linear fits estimated on the underlying microdata.

Figure A.4: Effects of Teacher Test Score Value-Added on Well-Being and Mental Health



(a) High School Well-Being



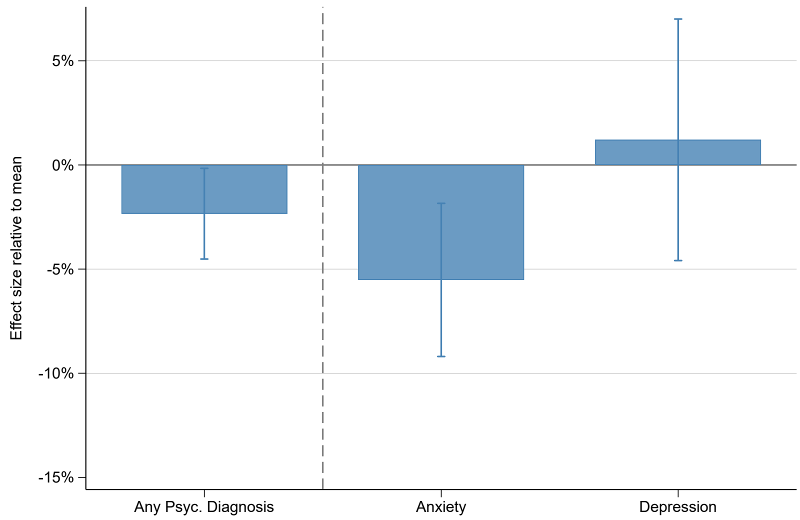


(b) Mental Health Care Utilization (Ages 16–18)

(c) Psychiatric Diagnosis (Ages 16–18)

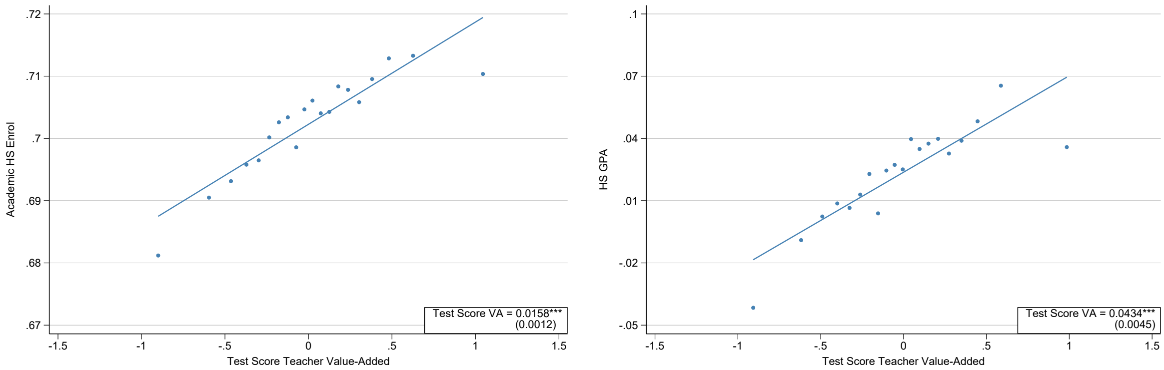
Notes: This figure plots the relationship between teacher test score value-added in middle school and students' well-being and mental health outcomes. In particular, we focus on three outcomes: (i) self-reported well-being during the first year of high school, (ii) indicator that takes value one if the student visited a psychiatrist or psychologist between ages 16 to 18 and (ii) indicator that takes value one if the student was diagnosed with any psychiatric condition between ages 16 to 18. The unit of observation is a student–subject–year. Both the outcome and the teacher value-added measure are residualized with respect to the baseline control vector used in the value-added estimation. Teacher value-added estimates are divided into 20 equal-sized bins, and the figure plots the mean residualized outcome within each bin against the mean value-added estimate, with the unconditional means added back. We also plot the best linear fit estimated before binning the data and report its slope coefficient and standard error, clustered at the student and classroom level.

Figure A.5: Effects of Test Score Value-Added on Psychiatric Diagnoses



Notes: This figure reports estimates of the effects of test score value-added on the probability of receiving specific psychiatric diagnoses between ages 16 and 18. The value-added measure is standardized and constructed using leave-year-out estimates. All regressions include the baseline control vector used in the value-added estimation. Standard errors are clustered at the student and classroom level.

Figure A.6: Effects of Teacher Test Score Value-Added on Academic Outcomes

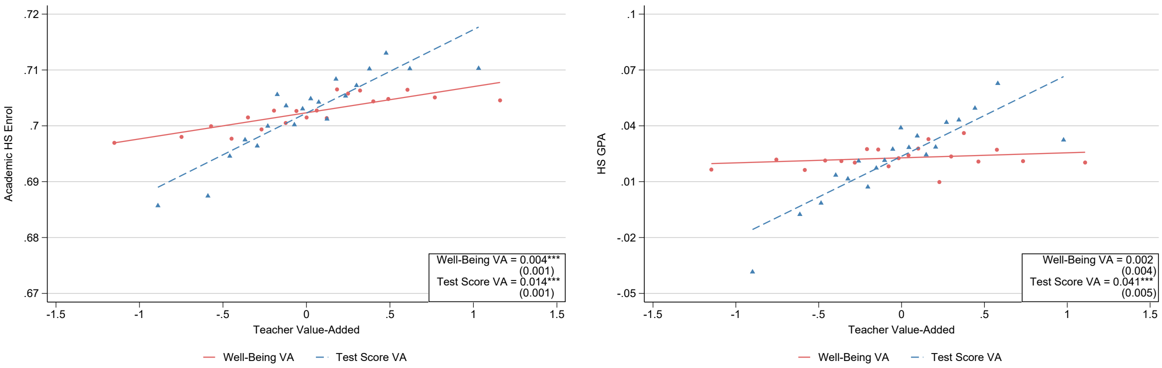


(a) Academic HS Enrollment

(b) High School GPA

Notes: This figure plots the relationship between teachers' test score value-added in middle school and students' academic outcomes. In particular, we focus on two outcomes: (i) indicator that takes value one if the student enrolled in academic high school and (ii) high school GPA, standardized by graduation year and academic track. The unit of observation is a student–subject–year. Both the outcome and the teacher value-added measure are residualized with respect to the baseline control vector used in the value-added estimation. Teacher value-added estimates are divided into 20 equal-sized bins, and the figure plots the mean residualized outcome within each bin against the mean value-added estimate, with the unconditional means added back. We also plot the best linear fit estimated before binning the data and report its slope coefficient and standard error, clustered at the student and classroom level.

Figure A.7: Effects of Teacher Value-Added on Academic Outcomes (Multivariate)

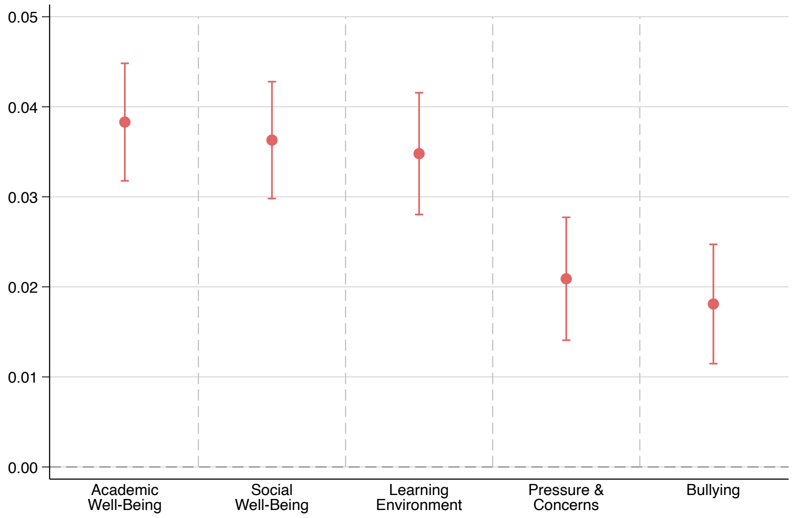


(a) Academic HS Enrollment

(b) High School GPA

Notes: This figure plots the partial relationship between teachers' value-added in middle school and students' mental health outcomes. IIn particular, we focus on two outcomes: (i) indicator that takes value one if the student enrolled in academic high school and (ii) high school GPA, standardized by graduation year and academic track. The figure displays two series corresponding to teacher well-being value-added and test score value-added, respectively; both series are based on a single multivariate specification in which the outcome is regressed simultaneously on both value-added measures. The unit of observation is a student–subject–year. Both the outcome and the teacher value-added measure are residualized with respect to the baseline control vector used in the value-added estimation. Teacher value-added estimates are divided into 20 equal-sized bins, and the figure plots the mean residualized outcome within each bin against the mean value-added estimate, with the unconditional means added back. We also plot the best linear fit estimated before binning the data and report its slope coefficient and standard error, clustered at the student and classroom level.

Figure A.8: Effects of Well-Being Value-Added on High School Well-Being Sub-Indices



Notes: This figure reports estimates of the effects of well-being value-added on the five high school well-being sub-indices constructed by the Ministry of Education: academic well-being, social well-being, learning environment, pressure and concerns, and bullying. The headline high school well-being measure used in the main analysis is the average of the academic and social well-being sub-indices. Each coefficient is from a separate regression of the standardized sub-index on well-being value-added, conditional on the baseline control vector used in the value-added estimation. Error bars indicate 95 percent confidence intervals, with standard errors clustered at the student and classroom level.

Table A.1: Quasi-Experimental Estimates: Effects of Test Score Value-Added on Long-Run Outcomes

	HS Well-Being	Psyc. Visit	Psyc. Diag.	HS Enrol	HS GPA
Test Score VA	0.0223***	-0.00518**	-0.00586***	0.0190***	0.0431***
	(2.58)	(-2.03)	(-3.03)	(7.00)	(4.16)
Observations	19,289	19,595	19,773	24,213	12,299
R-squared	0.001	0.001	0.001	0.046	0.001

Notes: This table exploits quasi-experimental variation in teacher quality generated by staffing changes across adjacent cohorts within school–grade cells. The unit of observation is a school–grade–cohort. Each column reports the effect of a one standard deviation increase in test score value-added on cohort-level high school outcomes, using within school–grade variation in teacher quality. Specifically, we regress changes in the school–grade–cohort mean of each outcome on changes in the corresponding school–grade–cohort mean of teacher value-added, estimated using leave-two-year-out measures of teacher quality. All specifications include year fixed effects and are weighted by cohort size. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Appendix B: Data Appendix

B.1 Well-Being Measure

The well-being measure is constructed from the following survey items:

Academic Well-Being

1. What do your teachers think about your progress in school?
2. Are you succeeding in learning what you want to learn at school?
3. How often can you find a solution to problems if you just try hard enough?
4. How often can you accomplish what you set out to do?
5. Can you concentrate during lessons?
6. I am doing well academically at school.
7. I am making good academic progress at school.
8. If I get distracted during class, I can quickly concentrate again.

Social Well-Being

1. Are you happy with your school?
2. Are you happy with your class?
3. Do you feel lonely?
4. I feel like I belong at my school.
5. I like breaks at school.
6. Most of the students in my class are friendly and helpful.
7. Other students accept me as I am.
8. Have you been bullied this school year?
9. Are you afraid of being laughed at?
10. How often do you feel safe at school?

B.2 Mental Health Outcomes

Mental Health Care Utilization Mental health care utilization is measured using information from the *Psychiatric Central Research Register* and the *Health Insurance Register*.

- *Psychological counseling:* Psychological counseling is identified using the Health Insurance Register. Services provided by private psychologists are identified using the physician specialty code 63. Mental health services provided by primary care physicians are identified using physician specialty code 80 in combination with reimbursement codes for mental health diagnostic tests and counseling (802147–802149, 804003, 804021–804027, 804050, 804063, 804106, 804116, 806000, 806100, 806101).

- *Psychiatric outpatient visits:* Visits to outpatient psychiatric clinics are identified in the Health Insurance Register using physician specialty codes 24 and 26. In addition, all outpatient visits recorded in the Psychiatric Central Research Register are included. This register captures visits to psychiatric departments in both public and private hospitals, including inpatient, outpatient, and emergency room contacts.

Psychiatric Diagnoses Among hospital-based psychiatric visits, we define an indicator for any psychiatric diagnosis based on ICD-10 diagnosis codes starting with the letter `F`. These diagnoses are identified using the Psychiatric Central Research Register.

Specific Diagnostic Categories We further construct indicators for specific types of psychiatric diagnoses. The following categories are used in the analysis:

- *Attention-Deficit/Hyperactivity Disorder (ADHD):* Defined using ICD-10 codes `F90.0-F90.9`
- *Depression:* Defined using ICD-10 codes `F32.0-F33.9`
- *Anxiety disorders:* Defined using ICD-10 codes `F40.0-F48.9`

All diagnosis indicators are coded such that an individual is classified as having a given diagnosis if they receive at least one corresponding hospital-based diagnosis during the relevant age window. In the main analysis, mental health outcomes are measured between ages 16 and 18.

B.3 High School Well-Being Measure

The high school well-being measure used as the main outcome in the analysis is the average of the Academic Well-Being and Social Well-Being sub-indices, which are the closest equivalent to the middle school well-being measure. The Ministry of Education constructs three additional sub-indices from the same survey—Learning Environment, Pressure & Concerns, and Bullying—which we use to disaggregate the long-run effects in Sections 4 and 6. All five sub-indices are listed below.

Academic Well-Being

1. I am doing well academically at school.
2. I actively participate during lessons.
3. I submit my assignments on time.
4. I keep working until tasks are completed.
5. I am prepared for lessons.
6. I like learning new things.
7. I am motivated by the teaching.
8. I do my schoolwork because I am interested in it.
9. I am making good academic progress at school.
10. If I get distracted during class, I can quickly concentrate again.
11. If something is difficult for me in class, I can take action myself to move forward.
12. How often can you accomplish what you set out to do?
13. How often can you find a solution to problems if you just try hard enough?
14. Are you succeeding in learning what you want to learn at school?

Social Well-Being

1. Are you happy with your class?
2. I am good at working with others.
3. I feel like I belong at my school.
4. I support and help my classmates.
5. I get along well with my classmates.
6. The atmosphere in my class is characterized by understanding and respect for each other.
7. In my class, it is good to be active during lessons.
8. I can get help and support from my classmates.
9. I wish I attended a different school.
10. I am in contact with classmates outside of school hours.
11. I feel bad at school.
12. Do you feel lonely at school?

Learning Environment

1. I get feedback from teachers that I can use to improve in my subjects.
2. I have influence on teaching.
3. Teachers ensure that students' ideas are used in teaching.

4. Teaching motivates me to learn new things.
5. The teachers provide academic help when I need it.
6. The teachers respect me.
7. Teachers coordinate times for assignment submissions.
8. There are good opportunities at school to get support and guidance.

Pressure & Concerns

1. I remain calm in stressful situations.
2. How often do you feel pressured at school?
3. How often do you feel pressured because of homework?
4. How often do you feel pressured because of grades?
5. How often do you feel pressured because of your own demands and expectations?

Bullying

1. Been called nasty names, made fun of, or teased in a hurtful way?
2. Been threatened to do things I didn't want to do?
3. Been subjected to unwanted psychological or physical attention?
4. Digital bullying?
5. Been subjected to unwanted sexual attention?
6. Witnessed students or staff being pressured to do things they didn't want to do?