5EV Technical Appendix

Holly Hart, Christopher Young, Alicia Chen, Naureen Kheraj, Elaine M. Allensworth

12/15/2021

Contents

Introduction

This document describes the methods and technical details used to produce the report entitled *5Essentials Survey in CPS: School Improvement and School Climate in High Poverty Schools*. This is the second of two quantitative reports concerning the revalidation of the *5Essentials* Survey. Appendix A in the first report, *Supporting School Improvement: Early Findings from a Re-examination of the 5Essentials Survey*, provides additional details on the data used, including survey response rates, survey measure descriptions and reliabilities, assignments of students to schools, and variable descriptions. Both revalidation reports use a common survey data set (and thus the details listed above are relevant to both).

Methods

Data

This study used data from the Chicago Public Schools (CPS) district. It is the third largest public school district in the country with 378,199 students and 654 schools in 2018-19. This study analyzed 499,537 elementary school students and 275,175 high school students in grades 1-12 across 548 elementary and 147 high schools from the 2011-12 through 2018-19 school years. Neighborhood schools, magnets, charters, selective enrollment, and vocational high schools were included. Analyses were conducted separately for

elementary and high schools. Elementary schools typically serve students in kindergarten through eighth-grade and high schools typically serve ninth- through twelfth-grade students. Schools that served irregular grades of students were classified according to whether more grades served were in the typical elementary or high school grade range (e.g., a school where seventh- through twelfth-grade students attend was considered a high school).

Survey data was collected from students and teachers who had taken the *5Essentials* Survey at a CPS school in the study years. All survey measures were aggregated to the school. Individuals were not linked longitudinally between years. All respondents were included to provide a complete picture of the school's organization. If fewer than six respondents at a school answered items in a survey measure, that measure was not scored for that school.

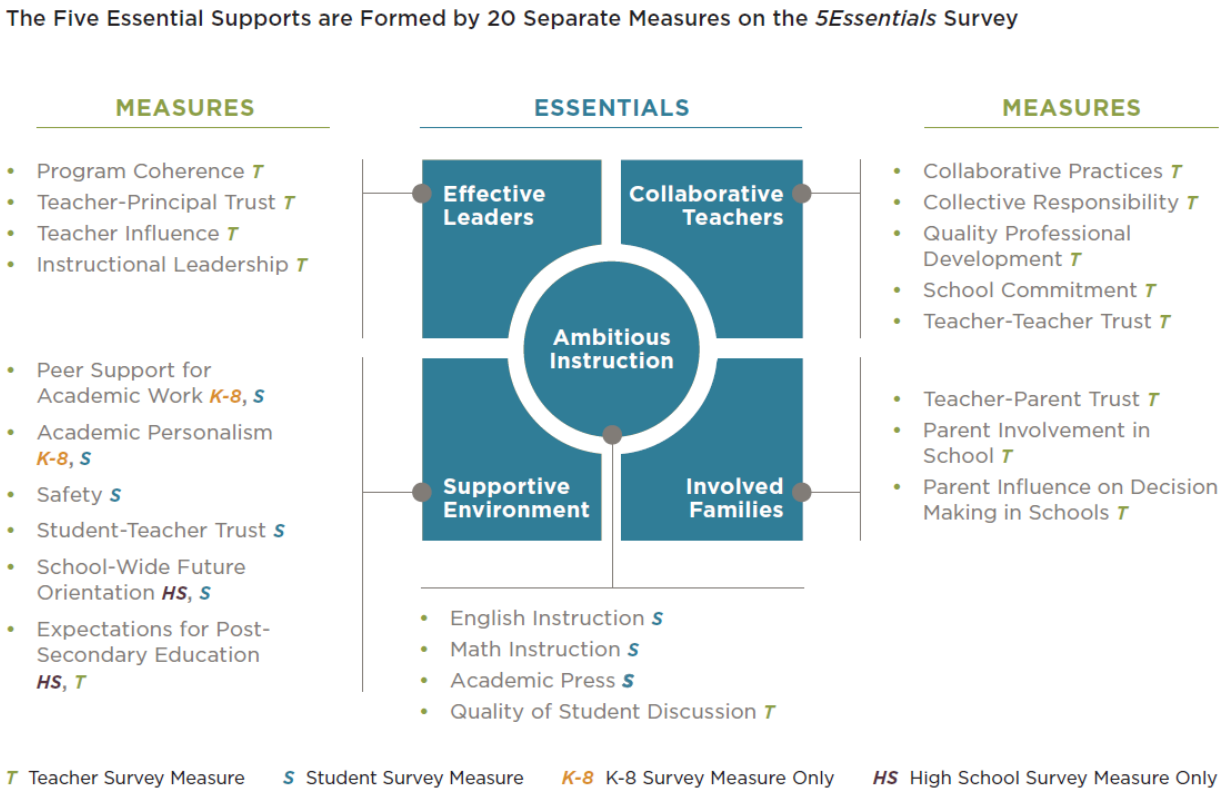
Students' outcome data was included in the analysis if they were in grades 1 through 12 in the year in question and had an active status in CPS's records. A student was only included in the estimate of their schools' outcome for a given year if they had attended their school for 45 days prior to the end of the spring marking period. Lastly, students were only included in outcome measures if they had attended schools in CPS for two consecutive years. These steps were taken to ensure that only students that had experienced the school environment continuously, and whose growth at a school could be compared to the prior year, were included in the estimate of the relationship of their school *5Essentials* rating with growth in their academic outcome.

Survey Measures

Data used to measure the five essential supports are based on annual student and teacher surveys (the *5Essentials* Survey) administered in the spring of 2011–19 to all schools in CPS. These surveys were administered online. Response rates for students in grades 6–12 range from 73 to 83 percent and for teachers between 65 and 83 percent.

Each of the five essential areas—Effective Leaders, Collaborative Teachers, Supportive Environment, Involved Families, Ambitious Instruction-–include four to five separate measures combined to create an essential score. Each measure is developed from responses to three to 10 questions on the survey.

Figure 1. *5Essential* Measures



Note: Measures that comprise the Supportive Environment essential are different for elementary schools (Safety; Student-Teacher Trust; Peer Support for Academic Work; Academic Personalism) and high schools (Safety; Student-Teacher Trust; School-Wide Future Orientation; Expectations for Post-Secondary Education). Thus, each *5Essentials* Survey includes 20 measures, but there are 22 unique measures.

Table 1. Measure Averages and Standard Deviations

Measure	Mean	SD
acno	3.42	1.28
colb	0.75	0.92
colr	4.60	1.77
engl	1.04	0.73
inf3	1.54	1.44
ins3	4.55	2.53
math	0.75	0.62
perc	3.70	1.19
pgmc	1.26	1.26
pnfl	1.44	1.95
pres	2.56	0.71
prt4	-0.08	1.65
qpd2	2.06	1.86
safe	2.07	0.93
scmt	3.53	2.15
sdis	3.77	1.21
slap	3.20	1.46
trpa	2.52	1.56
trpr	3.33	1.95
trte	5.76	2.06
trts	2.98	1.44
uexp	3.73	2.86

[1] Survey scores are presented in logits.

Responses to questions on the survey (survey items) are combined into measures through Rasch modelling (Rasch, 1993; Smith & Smith, 2004), using the BIGSTEPS program (Linacre, 2021). In the Rasch model, people and test items are placed on a single probabilistic scale.[1] Additionally, every item has defined 'step' difficulties, which estimate how difficult each relative response is (e.g., how likely a person is to "Strongly Agree" rather than "Agree" to that item). The item and step difficulties are used in the scoring of individual-level measures and producing standard errors in all survey years. Fit-inflated standard errors indicate whether the individual responded in a way that is well measured by the scale (e.g., responding with extreme responses to all items or endorsing 'hard' items but not 'easy' ones both lead to high error) and complete (e.g., responding to fewer items increases error). All survey measures were anchored using data from CPS teachers and students between 2014 and 2019. By anchoring the measures, researchers can make comparisons over time. Thus, a score on a measure will have the same meaning regardless of the year in which the survey was taken.

A school's score on a *5Essentials* Survey measure is calculated using precision-weighted means. First, the weighted average (also called "precision-weighted mean") of the individual measure scores was generated for each school. The weight is the precision (the inverse of the standard error). The measure scores of students with missing survey responses or extremely unusual response patterns are not very precise and are thus given less weight relative to the weight given to scores of other individuals when generating the school level score.

Outcome Measures

Student outcomes analyzed in this study were attendance, grades, and math test scores (English language arts test scores were highly correlated to math test scores and were excluded to avoid overweighting standardized tests). Attendance, grades, and test scores were obtained from CPS administrative data.

Test Score Data

Several different standardized tests were administered to CPS elementary and high school students in our study period. In the case of elementary school test scores, the ISAT, NWEA, and PARCC tests were administered during the study period, with multiple overlapping test scores given to the same students in several years. We used multiple imputation, a form of modelling where the relationship between scores is analyzed and used to predict missing scores, to put all elementary students' math test scores on the same scale. Specifically, we imputed all students' math scores onto the NWEA scale from all extant math test scores in a given year. For example, if in 2016 an eighth-grade student took both NWEA and PARCC math tests, we would impute both test results to create a novel NWEA score combining information from both test scores. In the case of high school test scores, students only had either an SAT or ACT score in any given year. All ACT scores were converted to SAT scores, according to a conversion table.

Table 2. Student Outcome Measure Averages and Standard Deviations

Level	Attendance_Mean	Attendance_SD	GPA_Mean	GPA_SD	Test_Mean	Test_SD
Elementary	0.95	0.05	3.13	0.66	218.82	20.90
High School	0.90	0.11	2.60	0.92	480.07	96.84

[1] Math test scores scaled to NWEA (elementary) or SAT (high school) scale.

All student outcome variables were standardized across the sample (so that "0" represents an average score across the entire span and "1" is a standard deviation above the average) prior to modelling.

[1]For example, when person score and item difficulty match, e.g., a person with a score of 1 responding to an item with difficulty 1, or a person with a score of 2 responding to an item with difficulty 2, the person has a 50 percent probability of endorsing that item. When person score and item difficulty differ the probability of endorsement is determined by a logistic scale, e.g., a person with a score of 0 has approximately a 73 percent probability of endorsing an item difficulty -1 and a 27 percent probability of endorsing an item of difficulty 1.

Additionally, the attendance and GPA data were transformed through the arcsin function to account for ceiling effects in these data (there was little variation for students at the top of the distribution before transformation).

Composite Creation

We aggregated data from multiple outcome measures for each student into a single composite value (Student Improvement Index) and also created a measure of each school's relative poverty based on census data standardized against a citywide average (School Poverty Index).

Student Improvement Index

The School Improvement Index captures the amount of year-to-year improvement in the following student outcomes: attendance, GPA, and standardized math test scores. This index captures whether or not students at a school improved beyond the level suggested by their academic performance in the previous school year. This reflects improvement, not simply their current level of academic performance. Improvement was estimated by comparing how much a student's score had actually improved compared to the typical trend for students like them.

Student Poverty Index

In the last phase of analysis we carried out, we measured the effect of a school's poverty status directly on improvement and whether survey measures relate to improvement in the same way at schools serving students from neighborhoods with different rates of poverty. The School Poverty Index describes the level of concentrated poverty in the school population, relative to other schools in CPS. The index incorporates two variables from the American Community Survey (ACS): the percentage of families with income below the poverty line and the percentage of adult males employed. The ACS data was linked to each student through the census block group in which their home was located. This student data was averaged at each school to create the School Poverty Index that measures the average level of poverty of students in the school. Schools were standardized across the CPS district, separately for elementary and high school levels.

Analysis of the Relationship Between the *5Essentials* and School Improvement

The relationship between each survey measure and the composite outcome was calculated using a hierarchical linear model. The model consisted of three levels; a student-outcome level, a school/school-year level, and a school level. The purpose of including each of these levels was to estimate the effect of the school and survey measures on overall student growth, as reflected in a composite score for each student. A first pass composite analysis was conducted, followed by a secondary analysis which included the effects of school-wide average poverty.

Data Format

At the observation level, data were tabulated in long format so that each entry represented a unique score for a student with a unique prior year score and a dummy variable indicating both the grade of the student and the type of outcome, with each entry identified by both a student ID number and school/school-year ID number. At the school-year level, survey measure scores for each school were associated with a school/school-year ID number as well as a school ID. Finally, at the school level, each school was indicated by its unique school ID number, and in the case of the context analysis, the school's average concentration of poverty across the study period.

Hierarchical Levels

At the student-outcome level, each student's outcomes are included as unique entries, indicated by a fixed effect of the outcome type and a grade fixed effect. This allows all outcomes to be estimated simultaneously, showing how much the school improved all outcomes when aggregated at level 2. At the second level, the school/school-year level, a random effect was included for both the survey measures baseline and growth, capturing the effect of the survey measures on student outcomes. At the third level, a random effect of the school, controlling for consistent growth or decline that occurs at each school relative to the district. The coefficients of interest are β_{01} and β_{02}, which describe how much a school is predicted to improve in their outcomes when they have higher *5Essentials* measure baseline or growth scores, respectively. In addition, for the context analysis, the average concentration of poverty at each school was included as a predictor of β_{00}, β_{01}, and β_{02}. Additionally, the coefficients that estimate the effect of concentration of poverty, γ_{001}, γ_{011},and γ_{021} are important for the context analysis. γ_{001} shows differences in the degree to which schools with different poverty rates tend to improve/decline on outcomes each year. γ_{011} and γ_{021} show differences in how much *5Essentials* measure base strength or growth, respectively, can predict outcome improvement/decline for schools at different poverty rates (e.g., a significant positive coefficient would suggest that in schools with higher concentrations of poverty, strength in the *5Essentials* measure will predict more growth than at a school with average concentration of poverty).

The formulas for regressing measures of student outcome improvement on school-wide organizational and contextual variables are shown below. The formula for improvement is shown for all three outcomes included in the composite at the elementary and high school level (attendance, grade point average, and test scores).

Student-Outcome Level

$$Outcome = \pi_0 + \pi_1 * OutcomeDummy + \pi_2 * GradeDummy$$

School and School-Year Level

$$\pi_0 = \beta_{00} + \beta_{01} * SurveyMeasureBaseline + \beta_{02} * SurveyMeasureGrowth + r$$

School Level

$$\beta_{00} = \gamma_{000} + \gamma_{001} * Avg.Scon + u$$
$$\beta_{01} = \gamma_{010} + \gamma_{011} * Avg.Scon$$
$$\beta_{02} = \gamma_{020} + \gamma_{021} * Avg.Scon$$

Analysis of the Probability of Being Strong on *5Essentials* Measures

We conducted an analysis of the likelihood that schools at, above, and below the average level of poverty in CPS would achieve a strong or very strong rating on each of the measures in the *5Essentials* Survey. Schools were first categorized as being between within one standard deviation of average in terms of the poverty status of their students, one standard deviation or more above the average, or one standard deviation or more below the average (average poverty, above average poverty, and below average, respectively). Schools were also categorized as "Strong" in an organizational measure if their precision weighted mean was at least 1 standard deviation above the mean (and categorized as "Not Strong" otherwise). We then analyzed the relationship of organizational strength and school poverty with multinomial logistic analyses, where strength was regressed on poverty status. The formula used across these analyses is shown below.

$$\log(\frac{Pr(Strong)}{Pr(NotStrong)}) = \beta * PovertyStatus$$

References

Linacre, J.M. (2021). *Winsteps® Rasch Measurement computer program* (Version 5.1.7)[Computer Software]. Portland, Oregon: Available from https://www.winsteps.com/

Rasch, G. (1993). *Probabilistic models for some intelligence and attainment tests.* Chicago, IL: MESA Press.

Smith, E.V., & Smith, R.M. (Eds.) (2004). *Introduction to Rasch measurement: Theory, models and applications.* London, UK: Jam Press.