Agreement between Student and Teacher Ratings of Students' Social and Emotional Skills

Ming Yin, Ph.D., Julia Szabo, M.Ed., and Erin Baumgartner, Ph.D.

Photo by Kenny Eliason on Unsplash

Research Brief

for the Houston Independent School District

January 2023

Research Brief

Agreement between Student and Teacher Ratings of Students' Social and Emotional Skills

The Study of Social and Emotional Skills (SSES) is an international effort led by the Organization for Economic Co-operation and Development (OECD) with the Houston Independent School District (HISD) serving as the only U.S. site. Over 6,000 10- and 15-year-old students from 119 HISD campuses were selected to participate in SSES in the fall of 2019. In addition, their selected teachers were surveyed to complement student ratings and provide more information on how students were perceived. **This brief examines to what extent teacher ratings of students' SE skills agree with or differ from student self-ratings.**

Social and emotional (SE) skills refer to the process by which children acquire and effectively apply the knowledge, attitudes, and skills necessary to understand and manage emotions, set and achieve positive goals, feel and show empathy for others, establish and maintain positive relationships, and make responsible decisions (OECD 2015). Examining commonalities and differences between students' self-ratings and teachers' ratings of students' SE skills could provide opportunities for teacher professional development and school curricular support around SE skill development. **This study found that about half of teachers' ratings matched with students' self-ratings. Compared to other SE skills, the lowest level of student-teacher rating agreement was identified in Trust and Stress Resistance, whereas the highest level of agreement was found in Cooperation. After accounting for a variety of student, teacher and school characteristics, age cohort and course grades were the factors that more consistently help explain agreement in student-teacher rating.**

Of the 45 questions teachers and students were both asked about a student's SE skills, about half of the teachers' ratings matched with students' self-ratings. Rating agreement varied across skills:

- **Trust and Stress Resistance had the lowest levels of rating agreement compared to all other SE skills.**

- **Cooperation had the highest level of rating agreement in both age cohorts.**

Student-teacher rating agreement varied by age group, with 15-year-old students being more likely to match their teachers' ratings than 10-year-old students.

Students with higher grades were more likely to have ratings that matched their teacher's.

Introduction

Background

The Study on Social and Emotional Skills evaluated students' social and emotional skills via students' self-reports combined with teachers' reports of the student[1]. Theoretically, combining information from different raters to evaluate the same characteristics is one way to overcome the inherent limitations of each of the individual assessment approaches (Kankaraš et al. 2019). Practically, teacher reports can complement student ratings and provide more information on how students are perceived, which may be useful to the district in developing resources for teachers and students. Understanding where students and teachers have misalignment in their ratings of students' SE skills provides opportunities for teacher professional development and school curricular support around SE skill development.

Research Questions

This brief addresses the following research question:

To what extent do teacher ratings of students' SE skills agree with or differ from students' self-ratings?

Data and Sample

Data for this study came from the SSES student and teacher survey collected in fall 2019. This survey was administered to a randomly-selected group of 10-year-old and 15-year-old students from 119 campuses in HISD (OECD, 2019). One teacher per student was also asked to fill out a survey linked to the participating student. The teacher survey gauged teacher feedback on student SE skills and also asked a series of questions re-

lating to themselves, their school, their role as a teacher, and the individual classes and subjects for which they were responsible. Over 1,300 teachers provided valid responses in evaluating students' SE skills and answering questions regarding pedagogy, school climate and training opportunities. For comparison purposes, the same items were used in both student and teacher survey, although the number of items per skill varied depending on the respondents. For students, each skill consisted of eight items while teachers' reports had three items per skill. Though over 6,400 students participated in the study, about 4,900 students received ratings from their teachers. See Appendix I for more information on students and items included for the analysis.

Additional data came from the 2019-20 Public Education Information Management System (PEIMS), HISD fall 2019 course grade data, and Texas Academic Performance Reports (TAPR) 2019. See Appendix II for more information on the students included in this study.

Measures of social and emotional skills

Survey questions answered by both students and teachers were included in this analysis. There were three questions relating to each of the following social and emotional skills. These skills include:

- Cooperation (living in harmony with others)

- Empathy (understanding and caring for others and their well-being)

- Trust (assuming that others generally have good intentions)

- Assertiveness (able to confidentially voice opinions, needs, and feelings)

- Energy (approaching daily life with energy, excitement and spontaneity)

---

1   The SSES study included a parent survey. However, the low parental response rate (18.5% of parents of 10-year-old student respondents participated) has led the OECD to recommend not using parental data. Source: OECD

- Sociability (able to approach others)

- Emotion Control (effective strategies for regulating temper, anger and irritation)

- Optimism (positive expectations for self and life)

- Stress Resistance (effectiveness in modulating anxiety and able to calmly solve problems)

- Creativity (generating novel ways to do or think about things)

- Curiosity (interest in ideas and love of learning and intellectual exploration)

- Tolerance (open to different points of view, values diversity)

- Persistence (persevere in tasks and activities until they got done)

- Responsibility (able to honor commitments, and be punctual and reliable)

- Self-Control (able to avoid distractions and sudden impulses

Because of the number of questions, the conceptualization of each skill is different from the same skill indicated in other briefs of the series.

Student-teacher agreement

The original SE skills questions were on a 1-5 Likert scale. This study collapsed the categories of the two disagreement and agreement statements (Figure 1). Student-teacher rating agreement was defined in the following three scenarios: (1) when both the student and teacher chose "neither agree or disagree"; (2) when both the student and teacher chose "agree" or "strongly agree"; or (3) when both the student and teacher chose "disagree" or "strongly disagree". All other scenarios were defined as rating disagreement.

Analytical strategies

This study first compared question-specific percentages of rating agreement and disagreement between student self-ratings and teacher ratings by each cohort. To understand the ways in which student, teacher and school factors were linked to rating agreement or disagreement, multilevel logistic regression was conducted on each skill. Important student demographic, academic, and socioeconomic characteristics, as well as school demographic information were accounted for in the analyses. Specifically, the following student characteristics were included: grade level, English learner (EL) status, special education status, gender, race/ethnicity, economic disadvantage status, and average reading and math grades in fall 2019. Teacher and school characteristics were taken into account in the analysis. Teachers' pedagogy, classroom climate and student-teacher relationship information were retrieved from teacher questionnaire (see Appendix for more information). Campus-level demographic characteristics, including the percentage of students who were economically disadvantaged, percentage of students who were English learners and percentage of students who were Black, were retrieved from TAPR 2019. Differences between campuses were also accounted for.

FIGURE 1 **Definition of Student-Teacher Rating Agreement**

Original Likert scale
Strongly disagree (1)
Disagree (2)
Neither agree nor disagree (3)
Agree (4)
Strongly disagree (5)

| If student reported	And teacher reported	Conclusion
Neither agree nor disagree (3)	Neither agree or disagree (3)	Agreement
Strongly disagree (1) or Disagree (2)	Strongly disagree (1) or Disagree (2)	Agreement
Agree (4) or Strongly agree (5)	Agree (4) or Strongly agree (5)	Agreement
All other scenarios are defined as mismatch or dis-agreement		

Key Findings

> **About half of the teachers' ratings matched with students' self-ratings, with some variation across skills.**

Across each SE skill, the average agreement between student and teacher ratings was around 50% (Figure 2 and Figure 3). In other words, students and teachers had similar ratings of student SE skills around half of the time. Question-specific percentages can be found in Appendix I.

Rating agreement varied across skills. Among 10-year-old students, higher percentages of rating agreement were found in survey questions related to Sociability, Cooperation, Optimism and Curiosity. Among 15-year-olds, the highest percentage of rating agreement was found in questions on Cooperation.

Questions regarding Trust and Stress Resistance had the lowest percentages of student-teacher rating agreement compared to all other skills. For Stress Resistance, about half of teachers' ratings were lower than student self-ratings, indicating that teachers tended to underestimate students' skill levels of Stress Resistance. In contrast, a larger percentage of teachers rated higher than students on Trust (in others), except that a larger share of 10-year-old students self-rated higher than their teachers on the second question of Trust. Generally, teachers tended to overestimate students' Trust skill levels. Directions of rating disagreement on all survey questions can be found in Table 2 in the Appendix I.

FIGURE 2 **Average agreement percentage of student-teacher rating by skill (10-year-olds)**
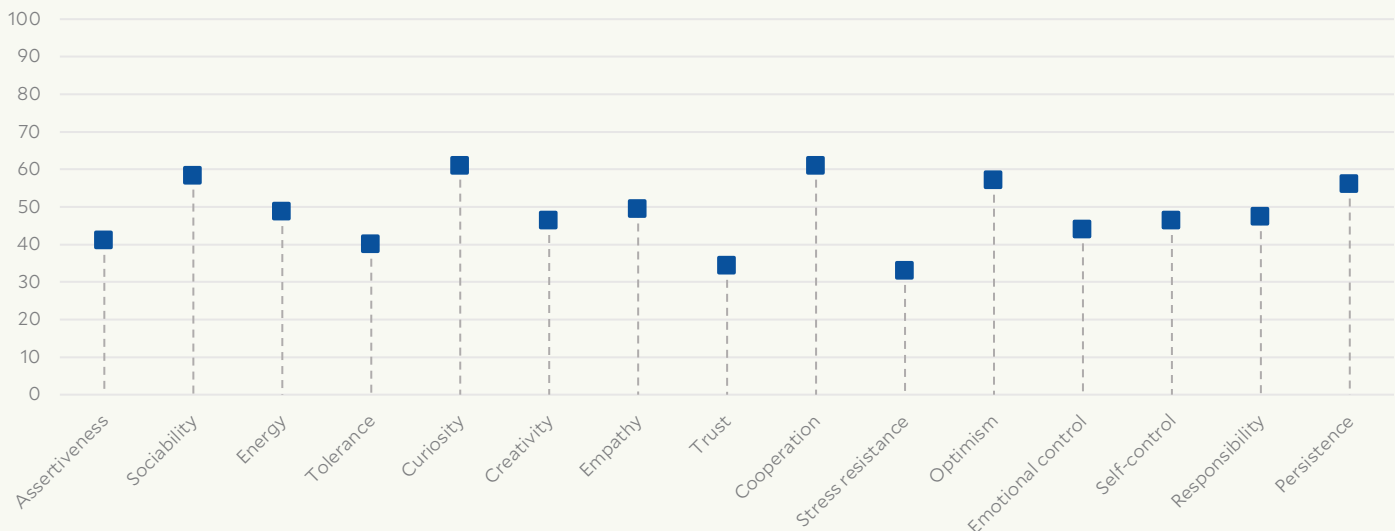
Note: Each skill demonstrated here included the three common questions answered by both students and teachers. The conceptualization of each skill might be slightly different from the same skill indicated in other briefs of the series.

Rice University | Houston Education Research Consortium

Average agreement percentage of student-teacher rating by skill (15-year-olds)
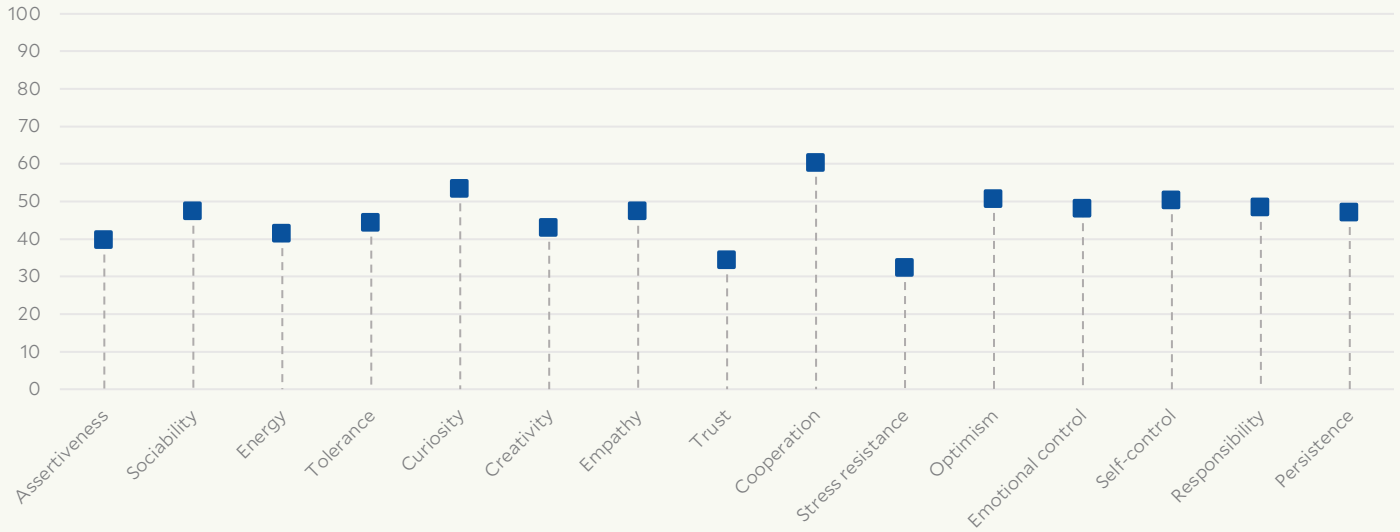
Note: Each skill demonstrated here included the three common questions answered by both students and teachers. The conceptualization of each skill might be slightly different from the same skill indicated in other briefs of the series.

Student-teacher rating agreement varied by age group, with 15-year-old students being more likely to match their teachers' ratings than 10-year-old students.

Student-teacher agreement in rating of student SE skills was higher among the older group (15-year-old students and their teachers), than younger group (10-year-old students and their teachers).

Students with higher grades were more likely to have ratings that matched their teachers'.

Course grades were also connected to student-teacher rating agreement. Students obtaining higher grades were more likely to have ratings that matched with their teachers.

Photo by Jeswin Thomas on Unsplash

Conclusion and Recommendation

Conclusion

The Study on Social and Emotional Skills (SSES) utilized multiple sources of information (i.e., student, teacher) to evaluate students' social and emotional skills. Combining student self- and teacher-ratings overall provides complementary information, contributing to more comprehensive understanding of student SE skills. Taking teachers' reports of student SE skills into account can also identify and correct for issues including self-perception biases and student-teacher disconnection (Kankaraš et al. 2019).

This brief compared the student self-ratings and teachers' rating of students' SE skills. Within the questions that were answered by both teachers and students, about half of the teachers' ratings matched with students' self-ratings. For both age cohort students, the lowest level of agreement was found in the questions about the skills of Trust and Stress Resistance. On the other hand, questions related to Sociability, Curiosity, Cooperation and Optimism for 10-year-olds and Cooperation for 15-year-olds exhibited higher levels of student-teacher rating agreement. This study also found age cohort and course grades were connected to student-teacher rating agreement. Fifteen-year-old students and students with higher course grades were more likely to have ratings that matched their teachers'.

Although both self-rating and others' ratings have their advantages and drawbacks, past research indicates subjective states such as personal experience, internal feelings and thoughts, are more difficult to perceive by external observers or raters (see Kankaraš et al. 2019). Resonating with this, the study found that skills including Trust and Stress Resistance had the lowest level of student-teacher rating agreement among all SE skills. Students demonstrating higher levels of Trust assume that others generally have good intentions and forgive those who have done wrong. Stress Resistance is a measure of students' effectiveness in modulating anxiety and ability to calmly solve problems. These two skills reflect a student's internalized characteristics. In contrast, Sociability (initiating and maintaining social connections), Curiosity (interested in ideas and love learning), Optimism (positive about life and self) and Cooperation (interconnected with others) are skills with more observable, externalized characteristics, therefore contributing to higher levels of student-teacher rating agreement.

Moreover, this study found age cohort and course grades mattered in predicting student-teacher rating agreement. Without proper instructions, younger students might be more likely than older students to lack skill-related knowledge, misinterpret the survey questions, give socially desirable answers, be inconsistent in answers or have memory biases (Kankaraš et al. 2019). Compared to other students, students with higher course grades might feel more connected to schools and interact with teachers extensively; in turn, teachers may know these students better.

Recommendations

- Create resources for teachers to better understand students' self-perceptions, particularly on the characteristics of Trust and Stress Resistance.

- Broaden pathways for building relationships and connections between students and teachers.

- Help students develop in articulating/identifying their SE skills, particularly for younger students.

Appendix I

TABLE 1 **Student sample included in the analysis**

Student Group	10-year-old cohort	15-year-old cohort
N	2636	2310
Grade		
	Grade 3: 0.3%	Grade 9: 22%
	Grade 4: 11%	Grade 10: 70%
	Grade 5: 83%	Grade 11: 8%
	Grade 6: 5%	Grade 12: 0.4%
Gender		
Male	49%	49%
Female	51%	51%
Race/Ethnicity		
Asian/PI	3%	5%
Black	21%	23%
Hispanic	68%	65%
White	8%	7%
Economic Disadvantaged Status		
No	18%	23%
Yes	82%	77%
English Learner (EL) Status		
Not Current EL	59%	79%
Current EL	41%	21%
Special Education Status		
No	93%	95%
Yes	7%	5%

Analytical samples were finalized based on the following exclusions. Students were excluded if (1) the demographic information was not captured in Public Education Information Management System (PEIMS) 2019-20 data; (2) survey responses were invalid or missing; (3) students identified as Native American or multiracial; or (4) SE skills not evaluated by a teacher. Percentages were calculated using survey weights.

| TABLE | 2 | **Question-level comparison between student self-ratings and teacher ratings and direction of disagreement** |

Survey question asked of both student and teacher	10-year-old		15-year-old	
	% Student-teacher agreement	**Direction of disagreement** *On average, teacher rated student's skills…*	**% Student-teacher rating agreement**	**Direction of disagreement** *Teacher's rating vs student*
Assertiveness				
1. A leader	42	↓ lower	39	↓ lower
2. Enjoy leading others	41	↑ higher	39	↓ lower
3. Like to be a leader in [their] class	40	↑ higher …than students rated themselves	41	↓ lower
Energy				
1. Full of energy	51	↓ lower	42	↓ lower
2. **Not** less active than other people*	45	↓ lower	40	↓ lower
3. **Not** have less energy than [their] classmates*	50	↓ lower	42	↓ lower
Sociability				
1. Have many friends	60	↓ lower	42	↑ higher
2. Like to spend [their] free time with others	63	↑ higher	51	↓ lower
3. Make friends easily	52	↑ higher	49	↑ higher
Creativity				
1. Original, come up with new ideas	43	↓ lower	42	↓ lower
2. Sometimes find a solution other people don't see	46	↓ lower	42	↓ lower
3. Have a good imagination	50	↑ higher	45	↓ lower
Curiosity				
1. Like learning new things	67	↓ lower	59	↓ lower
2. Like learning*	59	↓ lower	54	↓ lower
3. Love learning new things in school	57	↑ higher	47	↑ higher
Tolerance				
1. Ask questions about other cultures	36	↓ lower	48	↓ lower
2. Like hearing about other cultures and religions	42	↓ lower	44	↓ lower
3. Love to learn about other countries and cultures	42	↓ lower	41	↓ lower
Cooperation				
1. Like to help others	68	↓ lower	56	↓ lower
2. **Not** start arguments with others*	48	↓ lower	53	↓ lower
3. Treat others with respect	67	↓ lower	72	↓ lower
Empathy				
1. Important to [them] that [their] friends are okay	66	↓ lower	56	↓ lower
2. Can sense how others feel	43	↑ higher	43	↓ lower
3. Understand what others want	39	↑ higher	43	↑ higher
Trust				
1. Think most of [their] classmates keep their promises	34	↑ higher	35	↑ higher
2. Believe that [their] friends will never betray [them]	31	↓ lower	36	↑ higher
3. Believe that most people are honest	38	↑ higher	32	↑ higher
Emotional Control				
1. Keep [their] emotions under control	55	↑ higher	61	↑ higher
2. **Not** get mad easily *	45	↓ lower	46	↓ lower
3. **Not** have unpredictable emotions and moods*	32	↓ lower	37	↓ lower
Optimism				
1. Always positive about the future	46	↓ lower	42	↓ lower
2. Enjoy life	65	↑ higher	57	↓ lower
3. A happy person	60	↑ higher	53	↑ higher
Stress Resistance				
1. **Not** worry about many things*	31	↓ lower	31	↓ lower
2. **Not** often feel nervous*	35	↓ lower	34	↓ lower
3. **Not** often worried about something*	32	↓ lower	32	↓ lower
Self-control				
1. Can control [their] actions	56	↑ higher	66	↓ lower
2. Think carefully before doing something	46	↓ lower	47	↓ lower
3. **Not** say the first thing that comes to [their] mind*	37	↓ lower	38	↓ lower
Responsibility				
1. Reliable and can always be counted on	53	↓ lower	54	↓ lower
2. Keep [their] promises	52	↓ lower	51	↓ lower
3. **Not** often forget to do things [they] promised*	37	↓ lower	40	↓ lower
Persistence				
1. Keep working on a task until it is finished	58	↓ lower	46	↓ lower
2. Make sure that [they] finish tasks	57	↓ lower	48	↓ lower
3. Finish what [they] start	53	↑ higher	47	↓ lower

* This item was originally negatively worded—describing a student not having or having low level of the skill. For ease of interpretation, the items have been reverse coded and the questions here are positively worded—reflecting a student having high level of the skill. Example: The originally worded question asked students whether they "worry about many things." To better reflect a higher level of Stress Resistance (the header for the set of items), it has been reworded (and scores recoded) to reflect that a student does "not worry about many things".

Appendix II. Technical Notes

Analytical Strategies

Multilevel logistic regression models were estimated to answer the research questions. At the student level, gender, race/ethnicity, economic disadvantaged status, English learner status, special education status, and course grades included in the analysis. The complex survey design and student-level final weights were accounted for in all analyses. School differences were also controlled for by adding a school-level random intercept. All 15 SE skills were analyzed in the model separately.

Survey Weights

All analyses in this study have accounted for survey weights in order to calculate appropriate estimates of sampling error and to make valid estimates and inferences about the population.

The final student weight indicates the relative contribution of that unit to the estimated outcomes of the survey. It is the product of a design or base weight and of one or many adjustment factors. The former is the inverse of the selection probability and the latter compensates for random non-response and other random occurrences that could, if not accounted for, introduce bias in the estimates. These design weights and adjustment factors are specific to each stage of the sample design (OECD 2021b).

Indicators from Teacher Questionnaire

Active learning pedagogies
Teachers were asked to indicate how often various active learning pedagogies were occurring in their lessons ("never or almost never", "some lessons", "many lessons" or "every lesson or almost every lesson"). For analyses and scaling, the response categories "never or almost never" and "some lessons" were combined, given relatively low frequencies in these categories. Teachers received higher scores on this scale if they indicated they applied these learning pedagogies more often (OECD, 2021b).

Teacher pedagogies
Teachers were asked to indicate the extent ("not at all", "to some extent", "quite a bit" or "a lot") to which they could apply various teaching pedagogies. The response categories "not at all" and "to some extent" were combined into one category. Teachers received higher scores on this scale if they indicated they were able to apply the teaching pedagogies to a greater extent (OECD, 2021b).

School climate—quality of relationships
Teachers were asked about the school environment by indicating their level of agreement ("strongly disagree", "disagree", "neither agree nor disagree", "agree" or "strongly agree") with various statements. The categories "strongly disagree" and "disagree" were combined into one category for analysis and scaling because of low frequencies in these categories. Teachers received higher scores on this scale if they indicated a more supportive school environment (OECD, 2021b).

TABLE 3   Survey Questions of Indicators from Teacher Questionnaire

Active learning pedagogies	Teacher pedagogies	School climate—quality of relationships
▪ Students are given opportunities to explain their ideas. ▪ A small group discussion between students takes place. ▪ A whole class discussion takes place in which I participate. ▪ I discuss questions that students ask. ▪ Students present something to the rest of the class. ▪ Students discuss materials from a textbook.	▪ Get students to believe they can do well in school work. ▪ Help my students to value learning. ▪ Control disruptive behaviors in the classroom. ▪ Motivate students who show low interest in school work. ▪ Make expectations about student behaviors clear. ▪ Help students think critically. ▪ Get students to follow classroom rules.	▪ In this school, teachers and students usually get on well with each other. ▪ Most teachers in this school believe that the students' well-being is important. ▪ Most teachers in this school are interested in what students have to say. ▪ If a student from this school needs extra assistance, the school provides it.

Academic Outcomes

The academic performance control variable was composite grades created by averaging students' reading and math grades in fall 2019. The course grades data came directly from HISD and are different from the official end-of-year course grades data that the Houston Education Research Consortium typically receives. Most secondary schools had a six-week grading cycle. For schools with six-week grading cycles, the researchers used the Cycle 2 grades, which covered early October to early November.

To generate consistent performance measures on reading and math, a series of rules were applied. The researchers first categorized all courses offered based on subject areas and course descriptions. Among over 600 courses offered, 116 language-arts-related courses (i.e., English language arts, Reading, Writing) were categorized under reading and 59 math related courses (i.e., General math, Algebra, Geometry, Calculus) were categorized under math. Then, the researchers identified applicable reading and math courses for each of the sampled students. If a student had two or more courses within the same category, the course with the highest grade was kept. About 10% of students in each cohort had letter grades for applicable courses. Next, letter grades were converted to numeric grades based on HISD conversion guidelines.[2]

---

2   Conversion guidelines were based on grading scale in the 2020-2021 HISD School Guidelines on pp. XV-7: https://www.houstonisd.org/site/default.aspx?PageType=3&ModuleInstanceID=228302&ViewID=C9E0416E-F0E7-4626-AA7B-C14D59F72F85&RenderLoc=0&FlexDataID=191356&PageID=31617&Comments=true

References

Kankaraš, M. & Suarez-Alvarez, J. (2019). "Assessment framework of the OECD Study on Social and Emotional Skills", OECD Education Working Papers, No. 207, OECD Publishing, Paris, https://doi.org/10.1787/5007adef-en.

OECD (Organization for Economic Cooperation and Development). (2015). Longitudinal Study of Skill Dynamics in Cities. ESP (Education and Social Progress) Paris: OECD.

OECD (Organization for Economic Cooperation and Development). (2019). OECD Study on Social and Emotional Skills: Site Project Manager Manual Main Study 2019. Paris: OECD.

OECD (Organization for Economic Cooperation and Development). (2021a). Survey on Social and Emotional Skills (SSES): Houston (United States). Paris: OECD.

OECD (Organization for Economic Cooperation and Development). (2021b). OECD Study on Social and Emotional Skills: Technical Report. Paris: OECD.

Mission
The Kinder Institute for Urban Research builds better cities and improves lives through data, research, engagement and action.

About
The Houston Education Research Consortium (HERC) is a research-practice partnership between the Kinder Institute for Urban Research and 11 Houston-area school districts. HERC aims to improve the connection between education research and decision making for the purpose of equalizing outcomes by race, ethnicity, economic status, and other factors associated with inequitable educational opportunities.

RICE UNIVERSITY
Kinder Institute for Urban Research