# Tier Placement Analysis of the Satchel Pulse Platform

Michael A. Cook, PhD
Steven M. Ross, PhD

October 2022

JOHNS HOPKINS
SCHOOL of EDUCATION

Center for Research and Reform in Education

# Contents

# Tier Placement Analysis of the Satchel Pulse Platform

In May 2021, the Center for Research and Reform in Education (CRRE) at Johns Hopkins University contracted with Satchel Pulse to conduct an evaluation of the Satchel Pulse Social Emotional Learning (SEL) platform. This study was designed with two specific aims: to document research support for the primary components of the Satchel Pulse platform, and to analyze and compare extant outcomes data from students in classes that used the Satchel Pulse platform. This report addresses the second research aim described here.

In brief, Satchel Pulse SEL is a platform that identifies students with social and emotional skill areas that need further support and development. The platform employs a universal screener that identifies students who are in need of more support across any of the five CASEL Core SEL Competency[1] areas. These areas include self-management, self-awareness, social awareness, relationship skills, and responsible decision making. Following the screener, Satchel Pulse then makes recommendations for MTSS/RTI tiers and provides resources to support students' specific skill development needs.

The present study was designed to address the following research question:

1. How accurate are Satchel Pulse SEL tier placements?
    a. Are tier placements using both teacher and student scores more accurate than a system that would use either student or teacher scores? If so, to what extent?

## Method

### *Research Design*

This evaluation involved the analysis of data from the administration of the Satchel Pulse SEL platform assessment to students and teachers in the 2021-22 school year. The Satchel Pulse SEL screening assessment consists of a series of Likert-type rating items relating to student perceptions of their SEL, specifically across the five CASEL core competencies, as well as teacher perceptions of their students' SEL. Descriptive analyses of student and teacher Satchel Pulse SEL ratings were conducted, as well as analyses of classification accuracy of tier assignments, based on Pulse screener recommendations that only used student or teacher screener scores.

[1] https://casel.org/fundamentals-of-sel/what-is-the-casel-framework/

## *Participants*

Participants in this evaluation include students and teachers who used the Satchel Pulse SEL screener in the 2021-22 school year. CRRE was provided a dataset for analysis in this evaluation. Tier placement accuracy analyses were conducted using screener assessment data from a total of 94,693 students in 194 different schools.

## *Measures*

Data sources for the current study included anonymized individual student and teacher responses to the Satchel Pulse SEL screener assessment, averaged across competencies, as well as Pulse-recommended tier recommendations and accepted tier placements, as determined by school administrators. Satchel also provided CRRE with anonymized school IDs and time stamps of screener completion. In addition, Satchel provided tier recommendations if the screener consisted of only one user type (i.e., student only or teacher only).

**Student screeners.** The Satchel Pulse SEL screener assessment consists of a series of Likert-type items asking students to rate themselves on a series of 25 sub-skills derived from the five core CASEL competencies. Each item is related to one of the 25 sub-skills. These scores are combined to produce 25 sub-skill scores, which in turn are used to create five Competency scores.

**Teacher screeners.** After students complete the Satchel Pulse SEL screeners, teachers are then invited to complete different versions of the screeners for each of their students. Since this process would necessitate a teacher having to complete one screener for each student in their class, teachers are asked to rate students at each CASEL competency level, rather than at the sub-skill level, although teachers are asked about subskill information for Tiers 2 and 3 students. This is an important difference in the student and teacher screeners; students are required to rate themselves at the sub-skill level, while teachers rate their students at the CASEL competency level. This is done because requesting teachers to rate every student would be prohibitively time consuming and potentially unnecessary if teachers are capable of rating students at the competency level. Satchel provides information relating to CASEL competency definitions, so teachers are able to rate their students more accurately.

After student and teacher screeners have been completed, competency scores are combined, with the lower of the student and teacher competency scores being retained as the rating for each of the five CASEL competencies. For example, if a student's reply to sub-skills relating to Relationship Skills was 7.4, while the teacher's rating of the student on Relationship Skills was a 5, then the combined competency score for Relationship Skills would be 5, since this was the lower of the two scores. This process is repeated for each of the five CASEL competencies. The average of these five combined CASEL competency scores becomes a student's average SEL score, which is

one of the main outcome variables in the upcoming analyses. This process of retaining the lower of the teacher and student scores is performed to lessen the chance that a potential area of SEL concern is overlooked in a particular student.

Once the overall SEL scores are calculated from the student and teacher screeners, tier placements are determined, with the most at-risk students recommended for Tier 3 (i.e., those with the lowest SEL scores), and students with the highest SEL scores being recommended for Tier 1 placement. It is important to consider that tier cutoffs are created at the school level, meaning that tier placement cutoffs used at one school may not be the same as those used at another school. School-level tier placement cutoffs are determined by the information provided by schools relating to the proportion of the student population they can allocate to each of the three tier placement levels. Once the screening window is closed, an administrator at each school will be provided with Satchel Pulse's recommended tier placements for every student and will be asked to accept or modify the tier placement for each student. This final step results in the Approved Tier placement for each student, which is another important outcome variable in analyses. It is important to note that a student's Approved Tier is assumed to be the "correct" tier recommendation for that student. It is also important to note that, in the case of missing data (from either the teacher or student), competency scores and overall SEL scores and tier recommendations are derived only from the non-missing data source.

### *Analytical Approach*

Descriptive statistics, such as means and standard deviations, were computed for overall SEL scores. A variety of tier placement agreement metrics were computed to examine a series of different tier placement approaches, in regard to their classification accuracy.

## Results

We begin by descriptively examining overall SEL scores for students from the 2021-22 school year, followed by results of classification accuracy analyses.

### *Tier Placement Accuracy*

In examining tier placement accuracy, it is important to consider that there are significant proportions of observations that are missing either student or teacher screener scores. Specifically, 36.34% of observations were missing a student screener score, while 25.37% of observations were missing teacher screener scores. Thus, we will begin by examining tier placement accuracy using observations with non-missing student and teacher screener scores, followed by similar analyses for observations with one of the screener scores missing, and conclude by examining classification accuracy for the entire sample, regardless of missingness.

We begin by examining tier placement accuracy of students for which both student and teacher scores were available. This resulted in a sample size of 36,354, or 38.4% of the observations in the master data file. Table 1 displays Satchel Pulse recommended versus approved tier recommendations for these students.

Table 1
*Classification accuracy, Recommended vs. Approved Tier Placement*

| Recommended | Approved Tier | | | |
|---|---|---|---|---|
| Tier | Tier 1 | Tier 2 | Tier 3 | Total |
| 1 | 27,534 | 435 | 18 | 27,987 |
| 2 | 404 | 5,641 | 141 | 6,186 |
| 3 | 52 | 149 | 1,980 | 2,181 |
| Total | 27,990 | 6,225 | 2,139 | 36,354 |

*Note.* Analysis consisted only of observations with non-missing teacher and student screener data.

Classification accuracy was generally very high across all three tier levels, with overall agreement between recommended and approved tiers of 96.7%. Agreement was highest on Tier 1 recommendations, with 98.38% of approved Tier 1 placement agreement with recommended placements. This was followed by 91.19% of approved Tier 2 placements agreeing with recommended placements, and 90.78% of approved Tier 3 placements agreeing with recommended placements. Of the disagreements between approved and recommended tier placements, a vast majority (1,129 of 1,199 or 94.16%) of disagreements were within one tier level. These findings give evidence of strong alignment between Satchel Pulse and school administrator SEL tier placement recommendations.

Next, we examined tier placement accuracy of students for other subsets of observations. Tables 2 and 3 display the Satchel Pulse recommended versus approved tier recommendations for all observations with non-missing student scores and for all observations with non-missing teacher scores, respectively.

Table 2
*Classification accuracy for non-missing student scores, Recommended vs. Approved Tier Placement*

| Recommended | Approved Tier | | | |
|---|---|---|---|---|
| Tier | Tier 1 | Tier 2 | Tier 3 | Total |
| 1 | 54,798 | 473 | 65 | 55,336 |
| 2 | 535 | 10,750 | 212 | 11,497 |
| 3 | 27 | 148 | 3,755 | 3,930 |
| Total | 55,360 | 11,371 | 4,032 | 70,763 |

*Note.* Analysis consisted only of observations with non-missing student data.

Classification accuracy was again very high across all three tier levels, with overall agreement between recommended and approved tiers of 97.94%. As with the first set of analyses, agreement was highest on Tier 1 recommendations, with 99.03% of approved tier placements agreeing with recommended placements. Interestingly, Tier 3 placements had the next highest level of agreement, with 95.55% of approved Tier 3 placements agreeing with recommended placements. This was then followed by Tier 2 placements, with 93.50% of approved Tier 2 placements agreeing with recommended tier placements. Of the disagreements between approved and recommended tier placements, a large majority (1,368 of 1,460 or 93.7%) of disagreements were within one tier level. In general, agreement levels were slightly higher when considering all observations with student scores than when limiting to observations containing both student and teacher scores.

Table 3
*Classification accuracy for non-missing teacher scores, Recommended vs. Approved Tier Placement*

| Recommended Tier | Approved Tier | | | |
|---|---|---|---|---|
| | Tier 1 | Tier 2 | Tier 3 | Total |
| 1 | 45,084 | 985 | 196 | 46,265 |
| 2 | 711 | 9,278 | 261 | 10,250 |
| 3 | 30 | 188 | 3,551 | 3,769 |
| Total | 45,825 | 10,451 | 4,008 | 60,284 |

*Note.* Analysis consisted only of observations with non-missing teacher data.

When considering all observations with non-missing teacher data, classification accuracy was again very high, although slightly lower than in the previous analyses. Overall agreement between recommended and approved tiers was 96.07%. As with the previous analyses, agreement was highest on Tier 1 recommendations, with 97.45% of approved tier placements agreeing with recommended placements. This was followed by Tier 3 placements, with 94.22% of approved Tier 3 placements agreeing with recommended placements, and then by Tier 2 placements, with 90.52% of approved Tier 2 placements agreeing with recommended placements. Of the disagreements between approved and recommended tier placements, most (2,145 of 2,371 or 90.47%) disagreements were again within one tier level. In general, the subset of all observations with teacher screener data tended to have slightly lower agreement than the subset of all observations with student screener data, or the subset with non-missing student and teacher data.

Finally, we considered the classification accuracy of all students in the sample, regardless of whether either of the student or teacher scores was missing. The results of this analysis are found in Table 4.

Table 4
*Classification accuracy for either non-missing student or non-missing teacher scores, Recommended vs. Approved Tier Placement*

| Recommended Tier | Approved Tier | | | |
|---|---|---|---|---|
| | Tier 1 | Tier 2 | Tier 3 | Total |
| 1 | 72,348 | 1,04 | 209 | 73,611 |
| 2 | 811 | 14,387 | 324 | 15,522 |
| 3 | 39 | 195 | 5,326 | 5,560 |
| Total | 73,198 | 15,636 | 5,859 | 94,693 |

*Note.* Analysis consisted of all observations.

When considering all observations, classification accuracy in terms of tier placement is still very high, with overall tier placement agreement of 97.22%. As in prior analyses, the highest level of placement agreement was in Tier 1, with 98.28% of approved tier placements agreeing with recommended tier placements. This was followed by Tier 3 placements, with 95.79% of approved placements agreeing with recommended placements, and Tier 2 placements, with 92.69% of approved placements agreeing with recommended placements. Of the disagreements between approved and recommended tier placements, most (2,384 of 2,634 or 90.58%) disagreements were again within one tier level. Across the entire sample, patterns of agreement between recommended and approved tier levels were again very similar to those found in the subsamples previously described. It is important to consider that, in any given school, there is no way to ensure that all students and teachers are completing screeners, but there is no reason to expect there to be systematic differences between students who do or do not complete the screener, so this agreement is a reasonable expectation of what would be expected, on average.

## *Comparison to Single-user Type Designs*

In this section, we overview similar agreement analyses that consider a version of the Pulse screener that uses a single-user design, meaning only student scores or only teacher scores. While it may seem trivial that a system that uses both student and teacher data should achieve more accurate tier placements of students, the extent to which this occurs is unknown. The purpose of the following analyses is to examine the extent to which the combined Pulse screener tier classification system is more accurate than a system that only considers data from one source.

**Student-only tier placements.** Next, we examine tier classification accuracy statistics considering only student scores. We will examine this by comparing approved tier placements with student tier placements recommended by a system that only used student data. Table 5 shows the results of this analysis.

Table 5

*Classification accuracy, Student vs. Approved Tier Placement*

| Approved Tier | Student Tier | | | |
|---|---|---|---|---|
| | Tier 1 | Tier 2 | Tier 3 | Total |
| 1 | 24,522 | 3,152 | 316 | 27,990 |
| 2 | 3,069 | 2,134 | 1,022 | 6,225 |
| 3 | 890 | 434 | 814 | 2,139 |
| Total | 28,481 | 5,720 | 2,153 | 36,354 |

*Note.* Analysis consisted of student-only screener data from observations with non-missing student and teacher data.

Classification accuracy is considerably lower when considering only student score tiers in relation to approved tier placements, with overall tier placement agreement of 75.57%. Similar to prior analyses, agreement was highest on Tier 1 recommendations, with 87.61% of approved Tier 1 recommendations agreeing with student tier assignments. Agreement rates were considerably lower for Tiers 2 and 3 students, with only 38.10% of approved Tier 3 recommendations agreeing with student tier assignments, and 34.28% of approved Tier 2 recommendations agreeing with student tier assignments. The most common disagreements consisted of student scores indicating Tier 2 placement, while being approved for Tier 1 placement, and students' scores indicating Tier 1 placement, while being approved for Tier 2 placement (over 3,000 cases of each). There were also over 1,000 instances (2.81%) of student scores indicating Tier 3 placement, while being approved for Tier 2 placement. Of the disagreements between student and approved tiers, most (7,677 of 8,883 or 86.42%) disagreements were within one tier level.

**Teacher-only tier placements.** We also examine tier classification statistics using only teacher screener scores. We will examine this by comparing approved tier placements with student tier placements recommended by a system that only used teacher data. Table 6 shows the results of this analysis.

Table 6
*Classification accuracy, Teacher vs. Approved Tier Placement*

| Approved Tier | Teacher Tier | | | |
|---|---|---|---|---|
| | Tier 1 | Tier 2 | Tier 3 | Total |
| 1 | 25,923 | 1,992 | 75 | 27,990 |
| 2 | 2,137 | 3,521 | 567 | 6,225 |
| 3 | 318 | 400 | 1,421 | 2,139 |
| Total | 28,481 | 5,720 | 2,153 | 36,354 |

*Note.* Analysis consisted of teacher-only screener data from observations with non-missing student and teacher data.

Compared to using only student scores, classification accuracy is somewhat higher when considering teacher score tiers in relation to approved tier placements,

with overall tier placement agreement of 84.9%. As in prior analyses, agreement was highest on Tier 1 recommendations, with 92.62% of approved Tier 1 recommendations agreeing with teacher tier assignments. Agreement rates were again considerably lower for Tiers 2 and 3 recommendations, with 66.43% of approved Tier 3 recommendations agreeing with teacher tier assignments, and 56.56% of approved Tier 2 recommendations agreeing with teacher tier assignments. As with student tier recommendations, the most common disagreements in this analysis were teacher Tier 2 recommendations with approved Tier 1 recommendations, and vice versa. Of the disagreements between teacher and approved tiers, most (5,096 of 5,489 or 92.84%) disagreements were again within one tier level. In short, teacher tier placements agreed with approved tier placements to a considerably higher degree than did student tier placements. However, agreement between teacher and approved tier placements was still somewhat lower than agreement between recommended (which uses both teacher and student scores) and approved tiers and was considerably lower for Tiers 2 and 3 placements.

Overall, the use of a combined screener system that incorporates both student and teacher data, which is the design of the current Pulse screener, shows considerably higher classification accuracy of tier placements in relation to single-user systems that use only student or teacher data. Student-only data provided overall tier placement accuracy of 75.57%, while teacher-only data provided overall tier placement accuracy of 84.90%, both of which are considerably lower than the 97.22% tier placement agreement demonstrated by the current Pulse screener that uses both data sources.

## *Screener Descriptives*

We also conducted descriptive analyses of student and teacher Satchel Pulse screener scores, as a manner of examining score trends among both groups of screener participants. Table 7 shows average Satchel Pulse screener scores for students and teachers. We limited this analysis to observations with non-missing student and teacher screener scores.

Table 7
*Student and Teacher screener scores (n = 36,354)*

| | Mean | SD |
|---|---|---|
| Students | 7.280 | 1.461 |
| Teachers | 6.920 | 1.873 |

*Note.* Screener scores range from 0-10.

On average, student screener scores were generally slightly higher than teacher scores. Standard deviations show that teacher scores showed slightly more variability than did student scores.

To more closely examine patterns of Satchel Pulse screener scores for both students and teachers, we created histograms of score distributions from both sources. Distributions of student and teacher screener scores are found in Figures 1 and 2, respectively. As in previous analyses, we are restricting these analyses to observations with non-missing student and teacher scores.

Figure 1
*Histogram of Pulse student screener scores (n = 36,354)*

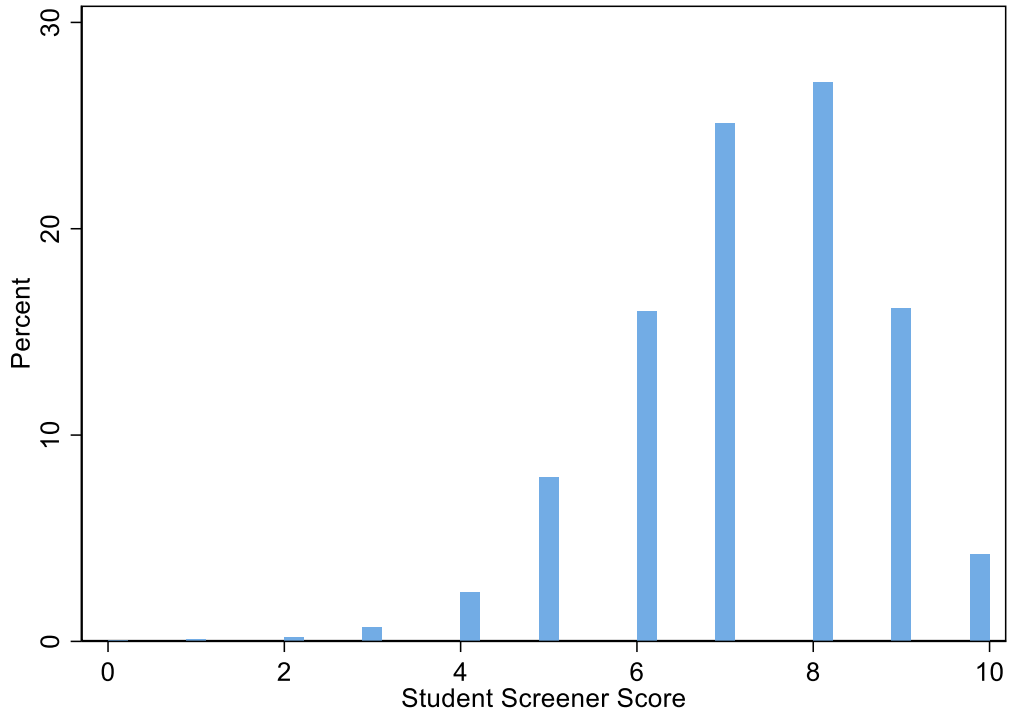
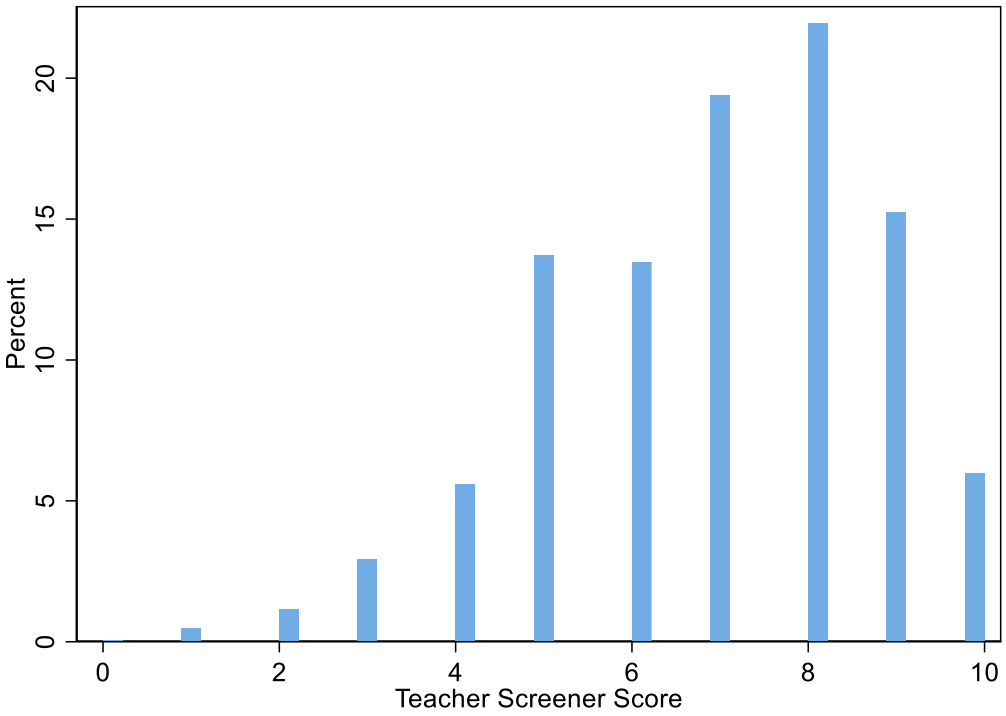



Figure 2
*Histogram of Pulse teacher screener scores (n = 36,354)*




Distributions of student and teacher screener scores are relatively similar, with the largest proportions of scores found for scores of 7 or 8 in both distributions. Student scores are more densely clustered around scores of 6-9, with 84.4% of student scores found in this range, while teacher scores are slightly more evenly distributed across the possible range of screener values, with only 70.1% of teacher screener scores in the 6-9 score range.

Overall, the Satchel Pulse SEL screener shows very high levels of classification accuracy, with overall classification accuracy between recommended and approved tier placements of nearly 97%. Importantly, classification accuracy was across all tier levels, with classification agreement percentages of over 90% for all three approved tier levels. These agreement indices were consistent when considering only observations with non-missing student and teacher screener scores, or when expanding to all student or all teacher screener scores and tier placements. We also found that the tier placement system that uses both student and teacher tier placements was considerably more accurate, in terms of agreement with approved tier placements, than was a system that used only student or only teacher screener scores. An initial descriptive analysis of student and teacher screener scores showed that student scores were slightly higher, on average, than were teacher screener scores, while teacher scores were somewhat more variable than were student scores.

## Discussion

This study was designed to investigate patterns of SEL scores and tier placements from a large database of students from nearly 200 schools in the 2021-22 school year. This report includes findings from a series of tier placement agreement analyses.

Results showed that SEL tier placements recommended by the Satchel Pulse platform agreed to a very high degree with tier placements approved by district administrators, with 97.22% of approved tier placements agreeing with tier placements recommended by the Satchel Pulse platform. Agreement was high across all three tier levels, with the highest agreement found among Tier 1 placements, followed by Tier 3 and Tier 2 placements. These levels of agreement were consistent across a variety of subsamples selected from the database, on the basis of missing and non-missing data sources. These findings give evidence supporting the use of the Satchel Pulse screener as a fast and efficient measure for reliably identifying students with SEL deficits and potentially in need of intervention.

Future research may include formal comparisons of Satchel Pulse screener scores from students who did or did not use Satchel SEL interventions or compare those who did or did not use the Satchel Pulse SEL screener on another SEL outcome measure.

The important conclusion of this evaluation is as follows:

- The Satchel Pulse SEL screener showed high levels of tier placement accuracy, with over 97.22% of recommended tier placements agreeing with school-administrator-approved SEL tiers.