Quantitative Analysis Technical Details:

Understanding statistical significance

When examining the results of the quantitative analyses described in this case study, the authors assessed the statistical significance of the results. Statistical significance communicates the likelihood of a study's statistical effect (for example, a difference between two timepoints or a relationship between two variables) appearing under a hypothetical scenario where there is no effect (in other words, the likelihood of finding an effect that does not actually exist). Typically, statistical significance is expressed in terms of a *p*-value; if the *p*-value is less than 0.05, the result is considered statistically significant. A *p*-value less than 0.05 means that we are likely to observe the effect less than 5% of the time in a world where there really is no effect. In other words, one would conclude that it is very unlikely that the finding is non-existent because the *p*-value is inconsistent with a hypothetical scenario of there being no effect.

Statistical significance is not a very intuitive concept and is often misinterpreted. Statistical significance does not communicate the size of an effect, how socially meaningful the effect is, or how likely a claimed effect is. Because the term *statistically significant* is often equated erroneously with the word significant in the everyday sense of the word (i.e. meaningful), this case study describes changes in values as *statistical differences*, rather than *statistically significant differences*. For example, the EASEL Lab at Harvard Graduate School of Education conducted paired t-tests to determine whether changes in SEL Benchmarks scores between the beginning and end of the summer program were statistically significant. Because the EASEL Lab has not yet determined a framework for determining whether a difference is *practically* significant, the authors of this report refer to *statistical increases* in SEL Benchmarks scores, rather than *statistically significant increases*. Another example is the BPS "deep dive," in which the authors analyze the relationship of child outcomes variables with time and measures of SEL implementation. When presenting the results of these analyses, the authors state whether each child outcome *changes statistically* over time (in other words, whether the association with time is statistically significant, whether or not it is practical significant), and whether each child outcome is *changes statistically* with SEL implementation values (in other words, whether the association with each SEL implementation indicator is statistically significant). In both cases, the authors identify statistical changes or associations to be those for which the p-value is less than 0.05.