# Instructions for EDI Data File (2015)

*Prepared by: UCLA Center for Healthier Children, Families, and Communities*

Enclosed please find:

1. Your 2015 EDI data in Excel format
2. A data dictionary.

You will be receiving your password in a separate email as the data file is password-protected. We hope you find this data file useful when examining relationships between observed variables and school readiness outcomes.

## Introduction to the data file

The data file contains the scored results of the Early Development Instrument (EDI) for children in your community, as well as demographic information on them. Each record in the data file corresponds to an individual child but, in order to protect student and teacher confidentiality, all identifying information has been removed. For example, we include child's age instead of date of birth, neighborhood instead of home address, and EDI ID instead of school ID. Records are scored based on the values associated with the answers given for each of the 103 items. The values of these individual items are not provided in the data file due to licensing agreements with Offord Centre. Instead, the data file includes the conversion of these items into variables which represent mean scores by domain and sub-domain along with four categorical variables indicating the child's level of development on each domain.

### Developmental Domains

In the data file, we provide the child's mean scores. Mean scores from the 2010 National Sample are used as a bench mark comparison to determine where the child falls on the developmental spectrum for each domain: Not On Track (Vulnerable or At Risk) and On Track (On Track-Middle and Very Ready). The developmental spectrum by domain is categorized as follows:

**Not On Track**: Child's domain score is less than or equal to the 25$^{th}$ percentile of the 2010 National Sample.

**Vulnerable**: Child's domain score is less than or equal to the 10$^{th}$ percentile of the 2010 National Sample.

**At Risk**: Child's domain score is greater than the 10$^{th}$ percentile and less than or equal to the 25$^{th}$ percentile of the 2010 National Sample.

**On Track:** Child's domain score is greater the 25th percentile of the 2010 National Sample.

**On Track – Middle:** Child's domain score is greater than the 25th percentile and less than the 75th percentile of the 2010 National Sample.

**Very Ready**: Child's domain score is greater than or equal to the 75th percentile of the 2010 National Sample.

### Developmental Sub-Domains

Based on a criterion scoring methodology developed by a team of experts, subdomains are categorized as "not ready," "somewhat ready," and "ready." This is different from how the domains are calculated. The subdomains are categorized based on a set values determined by developmental expectations.

### Additional Information

The data file also includes additional information that does not contribute to the EDI score, such as:

**Section A**: Demographic Information

**Section D**: Special Problems

**Section E**: Additional Questions

**Section F:** Customized district questions where applicable

Please refer to the data dictionary for more information about each variable.

## Data appropriate for analysis/reporting

EDI data is always reported on groups. The EDI is never reported on individual children or used as a screening or diagnostic tool for children.

For any analysis or reporting of scored EDI results, only valid records can be used. A value of one on the "valid" variable (valid=1) indicates that the record is valid for inclusion in analysis. When conducting analyses using only demographic information you need not limit your analysis to valid records only because scored results are not the subject of analysis, as described above.

Another important distinction to make when considering data for inclusion in analysis or reporting is sample size. To protect student confidentiality, the TECCS project policy maintains that licensed EDI recipients, including your team, **cannot** report or share data on any variable that has a category containing less than ten (10) records.

- If the analysis uses scored EDI results, a sample size of ten (10) or more valid records is needed for analysis.
- When reporting by geographic area, the neighborhood, block group, or zip code must have at least ten (10) valid records.
- If the analysis uses only demographic information, a sample size of five (5) or greater is needed. However, when your analysis includes multiple demographic variables, the minimum reporting value must be ten (10) or greater.

The EDI license agreement you received outlines appropriate use of data in much greater detail and all who use the data should be familiar with it.

## Overview of variables used to filter appropriate data

The following variables should be used to filter data when conducting analysis or reporting:

- *valid*: This means the data met the following criteria:
  - The child was in the classroom for at least one month
  - The child's teachers completed at least four of the five domains of the EDI.
- *School2015*: The analysis in the community profile is limited to records from the most recent year from which a school collected data. This variable indicates whether this record came from the most recent data collection for that particular school (school2015=1).
- *Suppress*: The "suppress" variable lets you know which neighborhoods have less than ten (10) valid records. A value of zero (0) on "suppress" means that the neighborhood is not suppressed and therefore can be used in reports because there are ten (10) or more valid records in that neighborhood. In contrast, a value of one (1) on the "suppress" variable means that EDI recipients **cannot** report data on the neighborhood since less than ten (10) children with valid EDI records live in that neighborhood. *In short, you should specify suppress=0 for any analysis run at the neighborhood-level.*
- *Count2015*: This variable gives the count of how many EDIs valid for analysis are included in each neighborhood.

## How to aggregate and disaggregate data for analysis/reporting

To disaggregate results by district, use either of the following variables:

- district_r: String variable (text) for district name
- district_id_r: Numeric variable for district.

In addition to the neighborhood names used in your Community Profile, we also provide the child's city, zip code, and census block group (obtained from the ArcGIS address locator used for geocoding all EDI records). Please refer to "data appropriate for analysis/reporting" above for important information on reporting data by geographic area.