

INDIVIDUAL-LEVEL DATA SOURCES

Data Acquisition and Sources

During the course of its ongoing research activities, the Juvenile Justice Educational Enhancement Program (JJEPP) obtains individual-level data from a number of sources each year. These data provide the basis from which to evaluate aggregate student performance in relation to various demographic and program characteristics and to assist in the specification of facility and student outcomes (i.e., credits earned, pupil progression, diplomas earned, return to school, and school attendance) and continuation of delinquency (e.g., arrest). The individual-level data used for research are obtained from the following sources:

- Department of Education (DOE) Survey 5
- Florida Department of Law Enforcement (FDLE)
- Florida Department of Corrections (FDOC)
- Florida Education and Training Placement Information Program (FETPIP)

The content of the submissions from each of these data sources is discussed below.

DOE Survey 5

Survey 5 contains a variety of reporting formats, but JJEPP's research initiatives are based on information contained in the following:

- Student Demographics
- Attendance
- Disciplinary Referral
- Dropout Prevention
- End-of-Year Status
- Exceptional Student Education (ESE) Status
- Transcript
- Entry/Exit Academic Assessment Testing

FDLE

FDLE is the source of arrest data for the measurements of the number of prior and post-release arrests as well as the severity of the crimes committed. A formal data sharing agreement was established with FDLE's Statistical Analysis Center (SAC). JJEPP supplies the SAC with a dataset of the FY2000-01 cohort, which contains offender identifiers, including: last name, first name, middle initial, sex, race, date of birth, and (when available) social security number (SSN).

Using these identifiers, the SAC matches the cohort to FDLE's Computerized Criminal History (CCH) database to extract all arrest records for any offender who is in both datasets. Only cases that match on an appropriate number and type of identifiers were retained as legitimate matches.

Arrest events with multiple charges are counted as one arrest. The types of arrest charges reported to FDLE include all arrests submitted by local law enforcement agencies through booking and fingerprinting in local jails and juvenile assessment centers in accordance with *Section 943.051, Florida Statutes (F.S.)*.

For additional information on the types of offenses and arrest information collected by FDLE please visit their Web site at <http://www.fdle.state.fl.us/>.

FDOC

FDOC data include all offenders' identification information and all sentencing events in its Offender-Based Information System (OBIS). To determine if, and when, the Florida Department of Juvenile Justice (DJJ) releases in the FY2000-01 cohort had been sentenced to prison subsequent to release, it is necessary to match the cohort cases to the FDOC offender identification information. The identifiers used included last name, first name, middle initial, date of birth, sex, race, and SSN. Various combinations of these identifiers are tested for matching accuracy; only in those cases where there is a high degree of confidence that the youth in the cohort is, in fact, the same offender in the FDOC data, is it determined that a valid match has been obtained.

For those cohort cases that match to the FDOC identification data, the FDOC offender identification number is used to match to the FDOC sentencing data to determine whether these youths had a prison sentencing date after their DJJ release dates. If so, the DJJ release dates are retained as part of the cohort data and used to create indicators to determine whether the youths had been sentenced to prison and the length of time from DJJ release to a prison commitment.

FETPIP

The data used to determine whether DJJ releases are employed is obtained from FETPIP. The SSNs of the FY2000-01 release cohort are shared with FETPIP as part of a data sharing agreement and were used to match to the quarterly employment data in their repository. Only employment records of those with SSNs that have been verified by the Social Security Administration are retained by FETPIP; therefore, if a youth provided an invalid SSN and was employed, there would be no match between the two datasets. Consequently, the number of employed youths reported may be an under representation of the actual number employed.

For those youths who have employment records, FETPIP supplied data on each year and quarter they were employed and their job code. The average salary earned during each quarter of employment was also part of the data FETPIP shared with JJEPP. In addition, military enlistment and college enrollment in four-year, two-year, and private Florida institutions are also obtained from FETPIP.

Creating a Cohort

Data for cohorts are selected using the school number from DOE Survey 5 data for a given year. Using the Master School ID list as well as the expertise of JJEPP staff, all juvenile justice residential commitment programs are identified by school number and selected from the Survey 5 Attendance File for FY 2000-01. This excludes any students who had earned diplomas prior to entering the juvenile justice program because they are not included in the

DOE data. Once identified, the cohorts were further reduced to only those youths who are released from their programs during the school year in question, based on withdrawal code and withdrawal date.

- Data obtained from DOE arrive in separate formats--Student Demographics, Attendance, Disciplinary Referral, End-of-Year Status, ESE Status, and Transcript--which must be *linked* together and later *matched* to other data sources, such as FDLE, FDOC, FETPIP, and JJEPP's own program-level quality assurance (QA) database.
- *Linking* within the DOE Survey 5 formats is done using student identification numbers (either SSN or an alias), school district, and school number.
- *Matching* to data sources outside DOE Survey 5 is done using SSNs and a combination of the students' names and dates of birth.
- Once data are grouped, linked, and matched, they may be summarized and analyzed.

Students may attend--and even be released from--more than one juvenile justice school within a given school year. In keeping with the notion of longitudinal follow-up, the *last* juvenile justice program from which the student was released is selected as the cohort record. Because follow-up analyses are calculated using release dates from the juvenile justice program, records with no release date are excluded. If a student's only DJJ record in the DOE Survey 5 demographic file is missing an exit date, that student cannot be retained in the cohort.

Widow and Orphan Records

Occasionally, data in the demographic format may not have a corresponding record in the attendance format. Conversely, a student who might otherwise be selected for inclusion in the cohort may have a line in the attendance file but not have a corresponding record at the same school in the demographic file. These records are called "widow" and "orphan" records. Widow and orphan records were excluded from the cohorts.

All records that contain a release date prior to the entry date into the cohort record program are discarded. All subsequent records are used for follow-up analyses.

A cohort file is then matched to subsequent years' of DOE data to build a placement history spanning the entire period from release to the end of follow-up in order to ascertain short- and long-term outcomes. Cohorts are further refined by examining student withdrawal codes after being linked to the Survey 5 attendance format and matched to subsequent years. Records that can be identified as "rollovers" (i.e., students who appeared in the same school the following year with less than a two-week break or who were only gone during the summer semester and did not have any other attendance record at a different school in between) are removed from the analyses because they had not actually been released during the school year. Withdrawal codes are also helpful in making determinations regarding releases; however, because many records did not contain a withdrawal code, it could not be the sole metric used in making the determination.

Tracking Student-Level Data Across Multiple Years

Only about two thirds of cases match from one year to the next in the Florida DOE Survey 5 demographic format.

Possible reasons why students may not be found in future Survey 5 data:

- Students obtained a high school diploma or its equivalent while committed to DJJ.
- Students may have left the state after their commitment.
- The local school district registrar never officially enrolled the student.
- The student's SSN or student identification (SID) number may have been reported incorrectly.
- The student died.
- The student dropped out of school.
- The student entered a private school.

Educational performance and outcomes are measured using the variables of return to school, arrest, attendance rate, employment, diplomas, credits earned, and pupil progression. Return to school is defined as whether a youth returned to a public, non-DJJ school following release. The measure of recidivism used is based on re-arrest and severity of offense using FDLE data.

Individual outcomes are also examined relative to the security levels of the program from which youths were released. The DJJ has a four-tier security and restrictiveness level system for its residential programs. In order of restrictiveness, the levels are as follows: low-risk residential, moderate-risk residential, high-risk residential, and maximum-risk residential/juvenile prisons. Day treatment programs often serve a mix of intensive probation, referral, prevention, and conditional release students. Because DOE student level data do not distinguish between these different types of youths served in day treatment programs, day treatment are excluded from the current cohorts.