Cradle-to-Career Definitions Subcommittee Meeting Summary

February 23, 2021

This document provides a summary of the key points that emerged from the February 2021 meeting of the Definitions Subcommittee. The suggestions from this group will be used to craft specific definitions for data points in the P20W data set. More information about the meeting, including support materials, a recording of the meeting, and the PowerPoint, are available at https://cadatasystem.wested.org/meeting-information/definitions-subcommittee (click on "Meeting Materials").

The goals of this meeting were to establish public display options for the following data points:

- Kindergarten readiness
- Transitional Kindergarten
- Kindergarten offerings
- Absenteeism
- Expected Days of Attendance
- Days of Attendance
- International Baccalaureate exams
- Military/Leadership Science courses
- Part time
- Debt

- Type of community college award
- Type of four-year institution degree
- Employment
- Earnings
- Living Wage
- High Demand Occupations

The following Definitions Subcommittee representatives attended the meeting:

Randy Tarnowski, Association of Independent California Colleges and Universities; Adrian Felix, California Student Aid Commission; Todd Hoig, California Community College Chancellor's Office; Erin Skubal, Marjorie Suckow, and Phi Phi Lau, California Commission on Teacher Credentialing; Channa Hewawickrama, Ben Allen & Paula Mishima, California Department of Education; Glen Forman, California Department of Industrial Relations; Janet Buehler, California Department of Technology; Matthew Case & Monica Malhotra, California State University; Angel Rodriguez & Margo Gonzales, Employment Development Department; Valerie Mendelsohn, West Coast University; Rima Mendez, California School Information Services; Chris Furgiele, University of California Office of the President

K-12 Data Points

Kindergarten Readiness

Channa Hewawickrama, Ben Allen & Paula Mishima of CDE clarified that the only available data point at the state level for kindergarten readiness is the Desired Results Developmental Profile score, which had been discussed at an earlier meeting and flagged as available by request only because it is only available for some students.

DECISION: The subcommittee categorized Kindergarten Readiness as "do not include."

Participated in a Transitional Kindergarten Program
There were no concerns about this data point.

DECISION: The subcommittee categorized Transitional Kindergarten as "ok to include."

Kindergarten Offerings
There were no concerns about this data point.

DECISION: The subcommittee categorized Kindergarten Offerings as "ok to include."

K-12 Chronic Absenteeism
There were no concerns about this data point.

DECISION: The subcommittee categorized K-12 Absenteeism as "ok to include."

Expected Days of K-12 Attendance
Paula Mishima of CDE explained that this data point is part of the calculation for Chronic Absenteeism. It is calculated on a per pupil basis, related to how many of the 180 days of the school year the student was enrolled in the California K-12 system. This data point might be useful in research studies but could cause confusion for those who are unaware of the calculated K-12 Chronic Absenteeism metric.

DECISION: The subcommittee categorized K-12 Days of Expected Attendance as "by request only."

Days of K-12 Attendance
The group concluded that, like Expected Days of Attendance, this data point would be most appropriate for research analyses rather than as an option on the query builder.

DECISION: The subcommittee categorized K-12 Days of Attendance as "by request only."

International Baccalaureate Exams
The subcommittee reviewed a list of options that included having a score of 4 or higher on one test or on two or more tests, to reflect the College/Career Readiness Indicator, which includes getting a 4 or higher on two or more International Baccalaureate tests. The CDE representative tentatively agreed to this framework and noted that the same approach should then be applied to Advanced Placement test scores, which were addressed at an earlier meeting. Jenny Singh, CDE's expert on the College/Career Indicator, will review the set of public display options that pertain to the College/Career Readiness Indicator and provide input on this approach.

DECISION: Pending further review by CDE.

Leadership/Military Science Courses
Similar to the International Baccalaureate tests, the College and Career Readiness Indicator sets a threshold level for leadership/military science courses. Therefore, the display options could include taking courses for less than a year, between one and two years, and two years or more.

DECISION: Pending further review by CDE.

Postsecondary Data Points
Part-Time Student
The group discussed alignment with federal definitions for part-time students, which provided the basis for the proposed display options. Similar to the Integrated Postsecondary Education Data System

(IPEDS), the proposed display options show enrollment on the basis of terms. However, the proposed display options would provide more specificity by looking at each term of enrollment, whereas federal reporting only references the first term that a student enrolls.

Todd Hoig of CCCCO suggested that the documentation should specify that only primary terms are included, such as Fall and Spring, as opposed to Winter intersessions or optional Summer enrollments.

Randy Tarnowski of AICCU suggested that as independent colleges with novel term formats, such as month-long options, join the Cradle-to-Career Data System, the display options should be reviewed to see how those specific term structures should be represented.

DECISION: The subcommittee categorized Part-Time Student as "ok to include."

Amount of Debt

The group discussed whether to represent debt as an average or a median. To inform this discussion, they reviewed the calculation used in the federal College Scorecard (the dictionary is available at https://collegescorecard.ed.gov/assets/CollegeScorecardDataDictionary.xlsx)

The College Scorecard calculates both median and average values, but uses only median on the public display. Chris Furgiuele of UC noted that averages can be calculated faster than medians and are more flexible analytically in cases where information needs to be regrouped to include different populations.

Randy Tarnowski of AICCU reminded the group that using averages means that a one or two outliers can skew the figures. He reflected on the inconsistency across data points in the federal display, which uses averages for earnings and medians for debt, and recommended that the methodologies be consistent across measures for the Cradle-to-Career Data System.

Chris Furgiuele of UC noted that it might be helpful for the query builder to include 25th and 75th percentiles for debt in addition to the median.

The group determined that both averages and medians should be made available in the query builder, and that 25th and 75th percentiles should also be considered for inclusion, but that medians should be used in the dashboard. However, this approach should be evaluated through the user centered design process.

Next, the group suggested adding a note to the documentation that clarifies debt should only be calculated for students who are no longer enrolled in college and should be assigned only to the institution from which a student graduated (as opposed to all colleges that the student attended, if they transferred between institutions before earning an award).

Finally, the group recommended that for items such as debt, employment, and earnings, the Cradle-to-Career Data System should align with definitions from the federal College Scorecard where possible.

DECISION: The subcommittee categorized Amount of Debt as "ok to include."

Type of Community College Award

Todd Hoig of CCCCO suggested that there may be times when the data points on Community College Award and Four-Year Institution Award should be combined. For example, bachelor's degrees conferred by community colleges should not be kept separate from bachelor's degrees conferred by four-year institutions.

DECISION: The subcommittee categorized Type of Community College Award as "ok to include."

Type of Four-Year Institution Award

The group recommended expanding the display options to separate the three types of doctoral degrees listed in IPEDS.

DECISION: The subcommittee categorized Type of Four-Year Institution Award as "ok to include."

Record of Employment

The group discussed the importance of clarifying the limitations of the state wage file, which will be the source of employment and earnings information for the Cradle-to-Career Data System. For example, those who work for the federal government, are employed out of state, are self-employed, or do not have a valid social security number are not included in the data set. This also means that public documentation should stress that the figure does not report the number of individuals who are unemployed. This data point will be most useful contextualize earnings data by indicating the relative sample size.

Given that each of the postsecondary partners calculates employment differently, and that there is no guidance at the state level or through IPEDS about how to do so, the group agreed that the federal College Scorecard provided a useful reference point.

The subcommittee contemplated the population for which a record of employment should be displayed, in the context of the dashboard specific to economic outcomes. They agreed that only those who completed college programs should be included (note: because CDE does not collect social security numbers, people who do not enroll in a postsecondary institution cannot be included).

Second, the group explored the timeframe in which employment records should be examined. The College Scorecard uses federal income tax records, and thus use calendar years for reporting employment and earnings. In contrast, the Cradle-to-Career Data System will use data reported by California employers on a quarterly basis. The group recommended using a calendar framework similar to the College Scorecard. So, for example, information displayed on students who graduated in the 2017-18 academic year would reflect employment information for 2019. Students would be considered to have a record of employment if they had earnings in any of the four quarters of 2019.

The group recommended convening a homework team to develop display options and underlying specifications for the employment data point, based on how earnings will be calculated.

DECISION: Pending further discussion by a homework team.

Earnings

The subcommittee discussed how to address students with less than four quarters of earnings data in a calendar year. Chris Furgiuele of UC noted that when students get a job part-way through the year, adding together earnings across the quarters will not represent that salary that they command at the end of the year. UC handles this problem by only including students with quarterly earnings records that also have earnings records in the prior and in subsequent quarter. UC applies this approach for the 20-25% of students who have fewer than four quarterly earnings records.

Todd Hoig of CCCCO noted that his agency uses different approaches in different contexts. For example, for the Student Centered Funding Formula, which must be calculated in a tight timeframe, earnings in the fourth quarter are multiplied by four to annualize them. However, for tools that are less time

sensitive like Salary Surfer, the Chancellor's Office adds quarterly earnings together. He cautioned against using data reported through the state wage file to project future earnings.

Chris Furgiuele of UC and Monica Malhotra of CSU wondered if the Cradle-to-Career Data System should only include students with four quarters of data.

Matthew Case of CSU clarified that different approaches would be appropriate depending on the use case. For example, if the intent is to counsel students on likely earnings for a specific major, a narrower data set might be more appropriate. But if the intent is to see empirically how much students earned, such as to clarify differential outcomes for different populations, then a more comprehensive data set would be better. He expressed a desire to work with EDD to identify additional data sources to provide information that may be missing from the state wage file.

Chris Furgiuele of UC concurred and added that if students are weighing majors or institutions, it will also be important to show earnings over time and not just entry-level wages.

Monica Malhotra of CSU noted that it would be important to exclude students who are still in college, for example students who have finished a bachelor's degree and immediately enrolled in graduate school. The federal College Scorecard only includes those who are no longer enrolled at any postsecondary institution.

Several members of the group noted that having the postsecondary segments work together to identify a common methodology for looking at earnings in various use cases would be a benefit for the state.

DECISION: Pending further discussion by a homework team.

Living Wage

Most of the data providers do not have a set methodology for calculating living wages. Matthew Case of CSU cautioned that using tools that are generated by external entities leaves the Cradle-to-Career Data System at risk of having to change data sources if the external entity ceases to provide the information.

DECISION: Pending further discussion by a homework team.

Location of Employment

Margo Gonzales of EDD noted that her agency is still researching whether a recent agreement to share employment records across states could be used for the Cradle-to-Career Data System. After the meeting she learned that the data sharing agreement does not allow the use of information from other states in the Cradle-to-Career Data System.

She suggested investigating whether it would be possible to provide information on where individuals are employed within California. However, after the meeting, she was informed that EDD only has information at the establishment level for employers, which would not be useful for understanding the actual location where an individual works. For example, an individual could work at a chain store in Redding that would be associated with a corporate headquarters in Los Angeles.

DECISION: The subcommittee categorized Location of Employment as "do not include."

Industry of Occupation

The group discussed the importance of helping the public understand the difference between industries—which describe the sector of employers, such as healthcare—and occupations—which

describe specific jobs, such as a nurse. This webpage provides more information about the data set, which could be used to craft these descriptions: https://www.census.gov/programs-surveys/economic-census/guidance/understanding-naics.html

Margo Gonzales of EDD noted that her agency publishes information by metropolitan statistical area at an additional level of granularity (3-digits) and that they track data at the 6-digit level. Matthew Case of CSU indicated that his agency displays a mix of 2-digit and 4-digit codes. In some cases, such as Healthcare and Social Assistance, two digits is sufficient, but in others, it is helpful to use 4-digit codes, such as for Elementary or Secondary Schools (a 4-digit code within the 2-digit code for Educational Services). In addition, it can be helpful to use higher-level groupings to ensure that there are enough students so that results do not have to be suppressed. Margo Gonzales underscored the necessity of masking data based on only a few individuals.

The group noted that EDD could provide information at the most granular level to the Cradle-to-Career Data System and the user-centered design process could be used to determine groupings are most useful once the deidentification protocol is applied.

DECISION: The subcommittee categorized Industry of Occupation as "ok to include."

High Demand Jobs and Earnings for High Demand Jobs
After clarifying how high demand occupations are calculated, the group discussed the differential between entry level jobs, which display hourly wages, versus middle skill and highly-skilled jobs, which are shown as annual earnings. Angel Rodriguez of EDD noted that annual earnings could be provided for all three categories of work. Matthew Case suggested that the methodology be explained to clarify any differences in how information is displayed in other contexts by EDD.

Next, the group discussed the 15 regional planning units used by EDD, which align with federal reporting requirements. Angel Rodriguez clarified that the regions had been constructed after a lengthy analysis, which is available here: https://www.edd.ca.gov/Jobs_and_Training//pubs/wsd20-01att1.pdf. Regional planning units are slightly different from the regions used by the community colleges, which was adopted by the subcommittee at an earlier meeting.

Given that labor market information will only be available using the regional planning unit groupings, some subcommittee felt that the Cradle-to-Career Data System should use this list instead, but others asked to check with colleagues on this question first.

Angel Rodriguez of EDD further clarified that wages are adjusted for inflation using the consumer price index. Given that earnings figures for students will need to be adjusted for inflation when displaying time trends, an EDD representative was requested to join the homework team to ensure methodologies align.

DECISION: The subcommittee categorized the following data points "ok to include:"

- Entry level high demand occupations
- Middle skills high demand occupations
- Highly-skilled high demand occupations
- Earnings for entry level high demand occupations
- Earnings for middle skills high demand occupations
- Earnings for highly-skilled high demand occupations

Additional Action Items

Given the concerns raised by several partner entities about the legality of sharing financial aid information, Kathy Booth of WestEd requested that agencies with concerns identify which statue is considered problematic, so that guidance can be solicited from the appropriate party.

For the next meeting, subcommittee members will be asked to review of a matrix of data points that have been approved by the subcommittee regarding three concepts: whether their agency will provide that specific data point, the context in which the data point can be accessed (for matching purposes only, by request only, in the query builder, and in a dashboard), and whether the information is about an individual or an institution. This matrix will be voted on at the March workgroup meeting.