Research Agenda Subcommittee Meeting Summary

March 13, 2020

This document provides a summary of the key points that emerged from substantive discussion over the course the day. More information about the meeting, including the background paper and the PowerPoint, are available at https://cadatasystem.wested.org/meeting-information/research-agenda-sub-committee.

The Research Agenda Subcommittee will identify parameters for research on six priority areas spelled out in the legislation. The March 2020 meeting had the following goals:

- Provide an update on the recommended scope for phase one of the California data system
- Ground the work of this committee by discussing user stories, use cases, and equity considerations
- Identify the possible scope of research and policy questions on long-term outcomes of primary school interventions
- Determine possible research projects
- List appropriate information that could be posted to public dashboards and query tools

The following representatives attended the meeting:

Tom Vu, Association of Independent California Colleges and Universities; Alma Mededovic, Bureau for Private Postsecondary Education; Tine Sloan, California Commission on Teacher Credentialing; John Hetts, California Community College Chancellor's Office; Ben Allen & Jonathan Isler, California Department of Education; Janet Buehler, California Department of Technology; Jennifer Schwartz & Chris Krawczyk, California Health and Human Services Agency; Martha Friedrich, California School Information Services; Jessica Moldoff, California Student Aid Commission; Dan Rounds, California Workforce Development Board; Tameka McGlawn, College and Career Academy Support Network; Muhammad Akhtar, Employment Development Department; Abraham Cicchetti, Gurnick Academy of Medical Arts; Lisa Lee, Hoss Lee Academy, Inc.; Alyssa Nguyen, RP Group; Steve Watkins, UnitekLearning.com; Tongshan Chan, University of California Office of the President; Jesse Rothstein, University of California, Berkeley; Michal Kurlaender & Sherrie Reed, University of California, Davis; Russ Rumberger, University of California, Santa Barbara.

Workgroup Update

The meeting opened with the facilitator providing an update on the topics covered in other subcommittee meetings during the month of February and March, as well as decisions made by the California Cradle-to-Career Workgroup at their February 26 meeting.

User Stories, Use Cases, and Equity Considerations

User Stories

First, the group discussed three possible user stories, which describe potential types of people who might access information from the state data system based on their level of comfort with quantitative data: data novices, data apprentices, and data experts (please see the meeting materials for full descriptions). Subcommittee participants noted that these designations are points on a continuum—for

example, individuals might fit between apprentice and expert levels. The participants offered suggestions for minor changes to the language that would make the user stories clearer.

P20W Data Set Use Case

Next, the facilitator summarized the draft P20W data set use case, which was provided in the meeting materials.

In examining the proposed list of data elements, the group discussed whether behavioral issues such as suspension should be included in the P20W data set. While one participant noted that behavioral issues can be helpful for analyses, others were concerned that there is bias in whether students are subject to disciplinary action. Given the sensitivity of this particular data element, it would be important to establish clear guidelines for when or how the information would be used. Another participant noted that it might be beneficial to use this data point in the context of looking at disparities in student experiences and the impact on longer-term outcomes. However, it would be important to ensure that data on disciplinary history would not be attached to specific individuals—such as in an electronic transcript. Another participant suggested that, in addition to only providing de-identified, aggregate data, the state could employ techniques that would make it difficult to reidentify information, such as grouping several data points together or providing information within ranges of values.

One participant flagged that the governance process should address cases where partner entities do not collect all of the data elements listed—would they be required to start collecting that information?

Other comments focused on the need to provide a more specific list of values, such as the types of disabilities that students have. Greater detail is also needed about the level of granularity that would be used in the contexts such as in a public dashboard versus for a research study. For example, dashboard might display only broad categories, whereas researchers could access more detailed data sets.

Another person noted that it will be important to determine which questions are most important for public dashboards. Other participants echoed the importance of having user experience testing to ensure the dashboards are well designed. Another noted that date novices, apprentices, and experts are likely to need different data points in order to answer their specific questions.

Several participants advocated for a mechanism that would allow users to upload a cohort of students and see their outcomes, in order to better understand the long-term impacts of specific interventions. One person suggested that this mechanism should also include a feature that generates a matched cohort for propensity score matching analyses.

Several ideas were posted in private chats to the facilitators but not discussed by the group, including:

- information on staff should be included in the P20W data set
- the concept of college readiness should be reframed as college access
- users should have the ability to save their preferences and come back to frequently used reports

Data Request Process Use Case

After reviewing the second draft use case, a participant noted that it would be important to set some basic parameters for the types of information that can be requested, so that partner entities do not

have to review requests that are not allowable by law. It will be important that the data host is well versed in federal and state regulations to assist with this task.

Another participant emphasized the importance of including an Institutional Review Board within the request process. The existing Committee for Protection for Human Subjects could be integrated into the state data system, but it would need to be prepared for the increase in requests.

For the partner entity data request process, one participant noted that it would be important to distinguish between data requested as part of an institution-based project, versus data being requested by an individual staff or faculty member for their own purposes, such as a dissertation.

The group also raised the question of whether the state data system would only be used to examine factors that involve only one agency or if two or more partners would be required for data access to be approved. Some argued that it would be preferable for the state data system to calculate all public-facing information to make it easier to find resources in a single place. Others expressed discomfort about the possibility of competing narratives between the state data system and individual agencies.

Operational Use Cases

For the draft operational use cases, one participant noted that there might be some concerns about adopting electronic transcripts because institutions rely on the fees that they charge to process paper transcripts.

Equity Considerations

The group discussed an article that outlined several ways that quantitative analyses can reinforce bias and identified implications for the types of information that should be included in various tools to help users identify interventions that might generate more equitable outcomes. One participant noted that many agencies do not collect the types of information that would help to identify structural challenges.

The group debated whether the data request review process could be structured to prioritize methodologies that investigate structural issues and to screen for bias. One participant argued that without an intentional focus on how data could be interpreted, inequities are likely to be reinforced. As one way to address this problem, another participant suggested that data requestors could be asked to specify the purpose of the analysis and how the results will be used. A third person noted that bias can surface in the conclusions drawn by the authors, even if the study design appears sound, which implies that there should be a requirement that reports are reviewed before they are released. Similarly, dashboards could be evaluated to ensure they don't reinforce narratives that blame students for poor outcomes without examining the student's context. However, one participant noted that if the state data system provides strict guidelines regarding what can be produced with the data, any study produced using the data set could be perceived as the official position of the state. The subcommittee recommended that the question of how results get vetted, both for analysis and in dashboards, should be examined as part of the governance process.

Long Term Outcomes of Primary School Interventions

After a presentation showing that there is little publicly-available data from other state data systems on the long-term outcomes of primary school interventions and outlining questions that researchers have

investigated using other data sources, the subcommittee brainstormed topics that would be most valuable to explore in California.

Topics fell into three general categories: education interventions, institutional context, and community context. Participants broke into small groups and dug into the initial list of topics to prioritize five questions, specify who would benefit from answers to each question, and how they would use this information (*note*: some small groups discussed topics assigned to other groups, so the notes below sort recommendations into the three main categories).

Education Interventions

Initial list of topics:

- Curriculum and standards
- Instructional delivery methods
- Learning outcomes (including academic and 21st Century skills)
- Course-taking and completion patterns
- Grade and course repetition
- Participation in ESL programs
- Participation in extracurricular/club activities
- Participation in college counseling
- Participation in special education programs

Priority questions	Who could act on this information?	How would they use it?
What impact does the timing and delivery of kindergarten have on college and career outcomes? What factors mediate this relationship?	- Parents - School officials - Policymakers - Early childhood educators - Researchers	Make more informed decisions about designing, funding, and participating in kindergarten programs
How does curricular differentiation and math placement in middle school impact college and career outcomes?	- Parents - School officials - Policymakers - Early childhood educators - Researchers	
What factors affect student needs for IEPs and 504 plans in primary school? How does participation in primary school IEPs and 504 plans impact secondary and post-secondary outcomes in subsequent years? How do the impacts vary among types of institutions (districts,		

cities, counties, charter/non-charter schools, higher education institution)?		
Which elementary school-level interventions have the greatest impact college and career outcomes? For specific interventions, are there differential student outcomes depending on the level of access, utilization, and implementation by the institution? What is the return on investment generated from the student outcomes for that specific intervention?	• Policymakers • Practitioners (teachers and staff, school leaders) • Parents • Researchers	• Understand how interventions are being implemented at each institution • Understand which interventions appear to have positive effects on student outcomes • Understand what components of interventions are yielding positive student outcomes • Understand which outcomes the interventions appear to be associated/influencing

Institutional Effects

Initial list of topics:

- Staff profile (experience, credentials, gender, race, age, turnover rate, salary differentials)
- Types of ongoing teacher training, professional development, and support
- Stability of leadership
- School resources and ADA funding
- Length of school day and school year
- School type (private, charter, public)

Priority questions	Who could act on this information?	How would they use it?
What kinds of training do teachers receive? What are the differential impacts of teacher training type (Teach for America, internship, type of higher education program, teacher certification/licensure status) on student outcomes over 10 years? 20 years? 50 years?	• State agencies • Practitioners (teachers and staff, school leaders) • Policymakers • Advocates • Community partners • Parents • Social services organizations (government and nonprofit)	• Understand how teacher training impacts student outcomes • Support requests for more investment in professional development such as resources, support, and teacher training
Are there matches/mismatches between student and teacher demographics and do they	• State agencies • Practitioners (teachers and staff, school leaders)	Understand how teacher and student demographic variances impact student outcomes

appear to affect student outcomes?		
What characteristics of the primary school (e.g., grade configuration, resources) impact college and career outcomes?	• Parents • Practitioners (teachers and staff, school leaders) • Policymakers • Researchers	Selecting schools
What are the effects of federal, state, and private funding streams on school resources? Within the context of these funding streams, does student achievement vary by race, ethnicity, socio-economic status? How do these effects impact student achievement throughout the student's life?		

Community Effects

Initial list of topics:

- Characteristics of students in the school (race, gender, socio-economic status, IEP/504 status, mobility, special ed status)
- Whether students are receiving social, food, and health services
- Parental employment, income, and savings
- Unemployment and poverty rate in the community
- Active military/veteran populations in the community

Priority questions	Who could act on this information?	How would they use it?
What are the long-term academic and workforce results of interventions on under-resourced communities?	• Social services organizations (government and nonprofit) • Parents • Legislators • Advocates	• Improve services • Request additional or different support • Change funding amounts and allocations • Better target programs
What interventions and investments can mitigate long-term negative impacts of social, food, and health disadvantages, particularly for students with high Adverse Childhood Events scores (ACEs)?	• School sites • Teachers • Social service leaders • Advocates • Families	Provide education and training to families, health and services providers, and teachers to provide mitigation and help provide services

How do regional unemployment rates and/or local economic characteristics affect student academic trajectories (course taking, postsecondary major, etc.)?	• School leaders • Policymakers • Politicians • Economic development boards	• Program development • Workforce pipeline development

How Data Should Be Shared

After sharing out on their discussions, the participants addressed the question of how these types of analyses could be shared—such as what would be appropriate for a dashboard versus what is better served through a research report. One person noted that because dashboards display descriptive data in graphical form, they are best for looking at trends or examining a variable such as gender or teacher preparation. However, dashboards don't tell the user why that variable might be predictive, because it cannot attribute causality or identify underlying factors.

To illustrate this problem, another participant used the example of teacher preparation. There are many different ways that teachers combine academic learning, ongoing professional development, and in-the-classroom experience. If a dashboard showed disaggregated student outcomes by type of teacher preparation, the results could be misleading. However, a dashboard could show the proportion of teachers with various types of training at each school, which might be valuable for planning purposes. But, the question remains, who would want to see this information? The California Department of Education (CDE) and parents might want different information about teacher preparation.

The group discussed the possibility of a rubric that could help determine whether information is appropriate for a dashboard or query tool. One participant thought that dashboards should only be used on topics that are well understood. Another participant noted it would be important to distinguish dependent variables (outcomes), which could be the data points on a dashboard, from independent variables (characteristics), which could be used to create disaggregation subgroups. A third participant suggested that the public dashboard should focus on a limited number of milestones and outcomes that have been prioritized by the state, while the query tool should provide information that helps to contextualize the dashboard metrics, and the firewalled analytical tools and research studies should enable a deeper investigation of the priority topics.

Another pointed out that many of the topics raised in the first two meetings of this subcommittee are already represented on dashboards that are hosted by the partner entities or federal agencies. It would be helpful to understand what information is already available and what is lacking, to inform what should be shown on dashboards for the California data system.

Next Steps

Thinking forward to the next topic—college readiness—participants suggested that the facilitation team identify how the concept is already being measured by various partner entities and identify how these data points could be displayed by the state data system. For example, a state dashboard could show the predictive K-12 factors found in the CDE School Dashboard as well as the outcomes of students after they enroll in postsecondary (paired with contextual information that might influence postsecondary

outcomes). Another participant clarified that it will be important to look at both college and career readiness, as this is the framework used by CDE.

The group also discussed how to drill down to a more granular research agenda. Participants suggested that it would be helpful to have homework assignments, such as taking individual research questions from the last two meetings, listing the data elements that would be needed, and identifying the best format for sharing the analysis in order to address specific ways that the data could be used. The homework can be done individually or in partnership with other subcommittee members (pairing up should be initiated by subcommittee members). Where possible, the subcommittee should work from research agenda questions from other states.