

## LESSONS LEARNED

# Working With Administrative Data in Adult Education



**Michelle Yin**  
**Stephanie Cronen**  
**Larry Condelli**

# Contents

|  | <b>Page</b> |
|--|-------------|
| Lessons Learned: Working With Administrative Data in Adult Education .....   | 1           |
| What Are Some Lessons Learned About Working With Administrative Data on Adult<br>Education Teachers and Students?..... | 1           |
| Recommendations on Data Collection.....  | 3           |
| Concluding Remarks.....  | 4           |
| References.....  | 6           |

## **Lessons Learned: Working With Administrative Data in Adult Education**

In program year 2010–11, more than 2 million students participated in adult education in the United States (National Reporting System [NRS], 2013b).<sup>1</sup> Participating adults and out-of-school youth come from diverse educational and linguistic backgrounds and have a wide range of goals and needs. Some rely on the adult education system to learn basic literacy or English skills, whereas others need help preparing for employment or preparing for high school equivalency testing and the transition to postsecondary education. Evidence from student outcomes suggests that adult education instruction helps students meet their goals; in 2010–11, for example, 48% of unemployed students who enrolled with the goal of finding employment—roughly 78,000 students—were employed shortly after leaving the program (NRS, 2013a).

Despite the important role that adult education can play in students' lives, research on how to promote adult education is limited. In the last few years, the availability of administrative databases in some states that track individual student achievement over time and link students to their teachers has brought new possibilities to the ways adult education programs and personnel can be evaluated. To provide descriptive information about the characteristics of teachers in adult education and to explore whether teacher characteristics are associated with student achievement in adult education, the Office of Career, Technical, and Adult Education (OCTAE) contracted with American Institutes for Research (AIR) to produce four research briefs.

The first three statistical briefs include a description of the characteristics of adult education teachers and examine teacher effectiveness, student achievement, and transitioning into postsecondary education. This brief—the fourth in the series—focuses on communicating common issues with administrative data and provides recommendations from a research and evaluation perspective.

### **What Are Some Lessons Learned About Working With Administrative Data on Adult Education Teachers and Students?**

Currently, the only large-scale source of data on adult education teachers and students comes from state data systems designed for federal (Workforce Investment Act) accountability reporting through NRS, which requires all states to have a student-level record system for reporting the outcomes, attendance, and characteristics of students who attend federally funded adult education and literacy programs. The quality of NRS data systems has improved over the years as advances in technology made data systems less expensive and more accessible. Likewise, the quality of the NRS data has improved as states gained more experience in collecting and reporting data. Consequently, a rich body of data exists among the states and local programs that can be used for secondary data analyses to answer research and policy questions.

---

<sup>1</sup> The U.S. Department of Education's Office of Career, Technical, and Adult Education provides funding to states for adult education through Title II of the Workforce Investment Act.

However, using NRS data for the purposes of analysis and research is not straightforward. Adult education data systems in most states are designed not for research but for annual state-level reporting to OCTAE. Also, the data systems often contain only NRS-required data elements, and the quality and subsequent usability of data vary across states. In the studies described in the three statistical briefs, AIR have requested student longitudinal data that would allow student-teacher matching from three states. As we cleaned and prepared the data sets for analysis, we noted the issues associated with using state NRS data for research and evaluation. Because states collect data primarily for accountability purposes and because the focus in the past has been on students and not also teachers, the following challenges are encountered when these types of data are used for research purposes:

- **Information on teachers is limited.** Before program year 2012–13, required data for teachers for NRS were limited to teaching status (full-time, part-time, or volunteer). As of 2012–13, data are also required on years of experience teaching adult education and on the type of certification held, which means that additional data on teachers will be collected in a uniform way at state and national levels. Currently, not all states collect information on teacher demographics, credentials, experience, and supports (e.g., professional development, paid preparation time). Having more data on teachers would allow a better understanding of how they contribute to adult education outcomes, which can be used by programs and states for program improvement purposes.
- **How information is collected is not consistent across states.** Each state has its own set of categories for data elements (e.g., teacher race/ethnicity, professional development, teacher and student educational attainment). Having more consistent definitions across states for data elements not only would be valuable to researchers but also would allow states to compare themselves with others.
- **It can be difficult to link students to their teacher.** Multiple teachers may be associated with the courses that adult education students take, perhaps as a result of coteaching. In addition, not all states have a unique identifier for each teacher. Having a clear link between students and their primary teacher is a fundamental requirement for certain types of analyses, such as analyses of teacher effectiveness.
- **It can be difficult to track teachers and students over time.** Although the three statistical briefs do not include longitudinal analyses, having data available over time (e.g., by using unique teacher and student identifiers) would allow researchers and program staff to better identify and understand effective teachers. As more adult education programs in states begin participating in Statewide Longitudinal Data Systems or developing other approaches to maintaining longitudinal data, this scenario may become a reality.
- **It can be difficult to determine when data are missing.** A basic need for data analysts is to be able to differentiate between missing information and a value of zero. For example, if no record is provided for a teacher’s professional development participation, analysts cannot determine whether the teacher’s data simply have not been entered or the teacher has not participated in any professional development. How these records are coded (e.g., left blank, coded with a “0,” coded with a missing data indicator such as “.”) can bias the conclusions drawn from the results if researchers do not accurately interpret the codes.

- **Self-reported data may bias findings.** Some data, such as teachers' participation in professional development, are collected through surveys rather than administrative records. These data may therefore lead to an overestimate or underestimate owing to teachers' difficulty in recalling the frequency and duration of their participation. Whenever possible, it is best to collect data from administrative sources.

## Recommendations on Data Collection

Given the issues identified, we provide the following recommendations that may help states maintain a data system that can be better used for their own analysis and program evaluation as well as for outside research.

- **Use consistent categories for teachers' and students' demographic data.** Currently, states collect data that are based on their individual needs and reporting purposes. There are no standard data categories at the federal level to guide the data collection process. For instance, some states categorize their teachers into seven racial categories (White, African American, Hispanic, Native American, Native Indian, Asian, Other), whereas others categorize teachers into four categories (White, African American, Hispanic, Other). For teacher and student education, the categories used are also not consistent within a state and across states. Having consistent categories is important not only for analytical purposes when each state evaluates its own teachers and students but also for comparing its students and teachers with those of other states on different measures.
- Collect standardized administrative data on teachers and include these data at the individual (nonaggregate) level in the state's data system. High-priority data not currently uniformly available; our recommended data values include the following:
  - Highest college degree held—no degree, associate's, bachelor's, master's, and doctorate or other professional (e.g., J.D.)
  - Certification—no certification, adult education certification, K–12 (or elementary/secondary) certification, special education certification, or TESOL certification; i.e., the categories recently added to NRS
  - Years of teaching experience in adult education as a continuous measure. (Ideally, other types of teaching experience would also be captured, such as teaching at the secondary-level.)
  - Hours of participation in program-related professional development—reported by program year (For the sake of standardized administrative data collection and tracking, this would include participation in events sponsored by the program and state, such as workshops or webinars on adult literacy instruction, but would not include activities that teachers choose to participate in on their own that cannot be easily recorded, such as taking courses at a community college that are not tracked by the program or pursuing a degree. Ideally, the hours would be broken up into topics; however, it would be difficult to standardize those topics across states and to apply those categories uniformly.)
  - Paid preparation time—this could be a simple yes or no variable (If the policy applies to all instructional staff in a program, the data can be entered at the program level.)

- **Create unique teacher identifiers to link student data to specific teachers.** Coteaching is popular in adult education, which presents a great hurdle for researchers who are evaluating teacher effectiveness. In addition, not all states have a unique identifier for each teacher that can be used to link to student data. If the state database cannot link individual teachers to students, it is impossible to relate teacher effectiveness directly to student outcomes or attendance. Consequently, policy makers and researchers cannot effectively evaluate the performance of individual teachers.
- **Improve state longitudinal data systems.** To examine teacher effectiveness over time, researchers need longitudinal data, which will allow them to follow the same students and teachers across years. There is a growing need to establish state longitudinal data systems for reporting and research purposes. Although the states that participated in our study possessed data systems that allow tracking students and teachers over time; we noticed inconsistencies when cleaning the data sets. For instance, states might not have a unique identifier for every student. When such students exit and reenter the program, they are treated as new students, which might bias analyses because they will be treated as a separate observation.
- **Have a separate category for missing data.** It is common in education data sets to have missing data. Many statistical methods are available to adjust for biases caused by missing data. However, if missing data are mistakenly treated as zero values in the data collection and analytical processes, biases might be introduced into the conclusions. Make sure the state's data system uses (and outputs) standard, nonzero values to represent missing data, such as a ".". Variables where this can be a common problem include hours of professional development, hours of attendance, student age, students' years of education, and teachers' years of teaching experience.
- **Avoid self-reported data.** Self-reported data have been shown to lead to biases in statistical analysis. The direction of biases depends on the variable. For instance, some states use student self-reported attendance hours to evaluate the relationship between attendance and performance. Students tend to overestimate their attendance hours, which might lead to upward bias when estimating its correlation with student achievement. The more reliable alternative is to record students' participation through a third party (e.g., teacher, program director) and combine information to calculate total attendance hours.

## Concluding Remarks

At the national level, our recommendation would be to increase guidance and encouragement to states for making the changes to data systems described above, either through existing mechanisms (e.g., NRS professional development opportunities) or through new support systems. Guidance should incorporate what is learned in K–12 research to the extent that it is applicable to adult education, but the incorporation should not rely solely on that research and should evolve to include what is learned from states and others who undertake research on teacher effectiveness going forward.

State staff and their approved researchers can then use the resulting data to answer their own policy or research questions, such as:

- What are the characteristics of adult education teachers in each state, and how do teachers vary within and across states? Do any meaningful patterns emerge that may be affected by the existing policies in each program or state?
- Do different types of students have equal access to qualified teachers?
- Do student outcomes vary by the types of teachers they have? What implications do these relationships have for policy and practice?
- Can models similar to those constructed for K–12 (e.g., value-added models) be used to reliably identify effective teachers in adult education? Would the use of these models be appropriate in the adult education setting, and if so, how should the models be constructed and used?

## References

National Reporting System. (2013a). *Core followup outcome achievement, program year 2010–11, all regions*. Washington, DC: U.S. Department of Education. Retrieved from <http://wdcrobcop01.ed.gov/CFAPPS/OVAE/NRS/reports/index.cfm>

National Reporting System. (2013b). *Participants by entering educational functioning level, ethnicity, and sex, program year 2010–11, all regions*. Washington, DC: U.S. Department of Education. Retrieved from <http://wdcrobcop01.ed.gov/CFAPPS/OVAE/NRS/reports/index.cfm>

## ABOUT AMERICAN INSTITUTES FOR RESEARCH

Established in 1946, with headquarters in Washington, D.C., American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally. As one of the largest behavioral and social science research organizations in the world, AIR is committed to empowering communities and institutions with innovative solutions to the most critical challenges in education, health, workforce, and international development.

## LOCATIONS

### Domestic

Washington, D.C.  
Atlanta, GA  
Austin, TX  
Baltimore, MD  
Cayce, SC  
Chapel Hill, NC  
Chicago, IL  
Columbus, OH  
Frederick, MD  
Honolulu, HI  
Indianapolis, IN  
Metairie, LA  
Naperville, IL  
New York, NY  
Rockville, MD  
Sacramento, CA  
San Mateo, CA  
Waltham, MA

### International

Egypt  
Honduras  
Ivory Coast  
Kyrgyzstan  
Liberia  
Tajikistan  
Zambia



AMERICAN INSTITUTES FOR RESEARCH®

1000 Thomas Jefferson Street NW

Washington, DC 20007-3835

202.403.5000

**[www.air.org](http://www.air.org)**

*Making Research Relevant*