New Ways to Ask Old Questions: Promising Avenues of Retention Research with State Staffing Data

Douglas B Larkin, Suzanne Poole Patzelt, and Liz Carletta

Montclair State University, Montclair NJ

Abstract: In this paper we describe the use of annual state-level school staffing reports as a data source for conducting research on teacher retention. Such staffing reports include salary, demographic information, educational attainment, and state certifications, and may be combined longitudinally to investigate questions related to teacher retention with an impressive scope and level of detail. Sample data from an ongoing National Science Foundation-funded project on teacher retention is shared, and demonstrates how such data may be used to identify cohorts of first-year teachers and track their persistence and mobility, including temporary exits from the workforce and subsequent returns. This paper concludes with suggestions for future research questions that could be investigated with the aid of these data.

Keywords: longitudinal data, science teacher education, teacher retention

One of the more pressing concerns of teacher preparation, mentoring, and induction efforts is the question of factors related to teacher retention. The focus of this paper is on the nature of the data used to answer such questions, and to suggest new possibilities based on an underutilized data source: annual state-level school staffing reports. Our findings come from a project on science teacher retention in which we have worked with a number of publicly available state-level school staffing reports. The impetus for this paper comes from a recognition that the data in these reports, when compiled longitudinally, permit education researchers to pose research questions in novel ways that may not have been easily answered using conventional data on teacher retention.

In the United States, much of the research on teacher retention has tended to draw upon two types of data sources. First are the large-scale surveys of teachers produced by the National Center for Education Statistics such as the Schools and Staffing Survey (SASS) and the Teacher Follow up Survey (TFS) used between 1987 and 2011. Data from the SASS have informed a great deal of foundational research in teacher retention research in the United States, particularly the work of Richard Ingersoll and colleagues (e.g. Ingersoll, 1997, 2007, 2011; Ingersoll & May, 2011,

2012; Ingersoll, Merrill, & May, 2016; Ingersoll & Smith, 2003). The successor to SASS, the National Teacher and Principal Survey (NTPS) has not been used yet to report on teacher retention, but rather to produce an annual report on the condition of education in the U.S. (McFarland, 2019). There are also smaller and more focused survey-based studies , such as the NJ Pathways study of a 1987 cohort over 11 years (Natriello & Zumwalt, 2017) and the later NYC Pathways study (Boyd et al., 2006).

The second type of data source informing teacher retention research comes from smaller-scale qualitative studies that track relatively small numbers of teachers longitudinally. For example, much of the literature around science teacher induction and mentoring has focused on tracking, describing or comparing different models of induction (Ceven McNally, 2016; Luft et al., 2011; Roehrig & Luft, 2006) or providing individual case studies of teacher learning during an induction program (Bang & Luft, 2014; McGinnis, Parker, & Graeber, 2004; Saka, Southerland, Kittleson, & Hutner, 2013). While these studies are valuable in understanding the particular experiences of novice teachers, they are somewhat limited in being able to inform policy. Given the often wide range of teacher education program quality (Zeichner, 2006) and variation in district and state mentoring and induction supports (Dawson, 2014), the ability to generalize from such in-depth studies may also be limited. While such studies are crucial in grappling with equity and justice issues in education (e.g. Achinstein & Barrett, 2004; Bianchini & Brenner, 2010; Lee, 2006), they may not point to salient trends in the teacher labor force that could meaningfully influence policymakers.

In this paper we describe a new kind of dataset that has reshaped the landscape for research in teacher retention: state-level school staffing reports. While in certain states these reports have been available for decades, the Race to the Top grant proposal process from 2009-2013 in the United States brought new attention to the pressing issue of the development of longitudinal data systems. As a consequence, many state-level education data systems now have unique teacher identifiers that allow for education researchers to examine questions about teacher retention (which includes teacher mobility, persistence, and attrition) at a scope and level of detail that was previously available only to state departments of education. Indeed, a growing number of researchers have gained access to these or similar state-level (or even large district-level) data to research teacher retention (e.g. Mandel, Fuller, & Pendola, 2018; Marinell & Coca, 2013; Simon & Johnson, 2015).

The development of these systems across states has been uneven (Boser, 2012; Flores, Park, Viano, & Coca, 2017), often focused more on student achievement data as opposed to directly examining teacher retention, but it is clear that soon many U.S. states will have the capacity to look at old questions about teacher retention in new ways.

The purpose of this paper is to sketch the possibilities for teacher retention research in this new data environment, and to influence the development of such systems to better facilitate knowledge generation about teacher retention across a variety of policy environments in different states. The specific research question to be examined here is: *What is the potential for using state-level school staffing*

databases for education researchers examining issues of teacher retention (which includes teacher mobility, persistence, and attrition)? Though the research project driving this inquiry is focused on secondary science teacher retention in high-need schools and on the retention of science teachers of color, the conclusions from this paper are broadly applicable to all demographic categories, grade-levels, and teacher certification areas.

Theoretical Framework & Data Sources

In our work, we draw heavily on the review by Borman and Dowling (2008) that examined the literature on teacher retention generally, and the review by Achinstein, Ogawa, Sexton, & Freitas (2010) that examined retention of teachers of color. Both provide a coherent theoretical framework for examining teacher retention, and classify factors into two broad categories: personal/professional and school contexts.

Personal/professional. Data related to personal and professional backgrounds includes certain common fields, such as first, middle, and last name, salary, and year of birth. Reporting of race and ethnicity have changed over the past decade, and given that states must report race and ethnicity data to the federal government, many state data systems appear to have adopted federal guidelines that allow for respondents to choose more than one race, and present ethnicity as a separate category (Spellings, 2007). Gender data is also included in this data set, and while some states have moved to include a non-binary response option for students

(e.g. Virginia and the District of Colombia), it is not clear that teacher-level data in any state currently includes this option.

In terms of professional data, these data systems include educational attainment level, state certification code(s), full/part time status, and perhaps most importantly , the school and local education agency (LEA) assignment.[1] Also typically included are years of experience in education in the state, in the LEA, and total experience. Some states have included a field for preparation pathway, which may be somewhat coarse (e.g. New Jersey data provides the option for the selection of "traditional" or "alternate route"). More specific data linked to the individual teacher like the name of the educator preparation program may also be collected, but is rarely publicly released.

School Contexts. When staffing data is used in combination with other district/school level data, it becomes possible to examine the relationship between retention and three of the school context categories (student body characteristics, financial capital, human capital), identified by Achinstein et al. (2010), which may include: starting salary, average district salary, district size, school size, and demographics of the school/district. It is also likely—when the sample size of teachers of color is larger—that it becomes possible to examine the social capital dimension from these data, which may include other factors such as proportion of

---

[1] The designation local education agency (LEA) is used in state data systems instead of "school district" so that the data may also include publicly available data from charter schools, which are not technically school districts.

teachers of color in the school, subject area department or grade-level size (a proxy measure for the number of teachers in the school with given certifications), as well as the match between the self-identification of the race/ethnicity of the teacher and the demographic profile of the school/district, which was noted as a factor related to teacher retention in the Zumwalt et al. longitudinal study (Zumwalt, Natriello, Randi, Rutter, & Sawyer, 2017; Zumwalt, Randi, Rutter, & Sawyer, 2017).

Findings

To illustrate the possibilities of using state-level data systems to examine issues related to teacher retention, we draw upon all of the Professional Personnel Individual Staff Data from 2007-8 to 2017-18 teacher level data from Pennsylvania, which is publicly available on the state Department of Education website.[2] In our research, we are specifically interested in teachers with a secondary science certification, and their mobility, persistence, and attrition during the first six years of teaching. By combining the data sets into a single database, selecting only teachers with a secondary science certification, and sorting the teachers into cohorts based on their (apparent) first-year of teaching, we are able to track how long teachers remain in their school assignments, when they leave and return to the workforce, and when they move from one school or LEA to another.

This process requires careful cleaning of the data, and there are marked differences in the data structures before and after 2013, when changes were made

[2] https://www.education.pa.gov/Data-and-Statistics/Pages/Professional-and-Support-Personnel.aspx#tab-1

to the state's educational data system in the wake of Pennsylvania's successful award of a $41 million Race to the Top grant. One notable difference is the addition of unique teacher identifiers in 2013, that were not present in the public-facing data, but were released without issue to the research team upon request. This identifier made it much easier to track teachers—particularly in cases where surname changes occurred. Year of birth, and race/ethnicity data also did not appear in the public data, but was similarly released by the Pennsylvania Department of Education upon request.

For illustrative purposes, we conducted two analyses of these data. In the first (see Figure 1) we looked at retention of science teachers in Pennsylvania public schools by first year cohorts. In the second (see Figure 2), we combined this teacher retention data with school-level population data to create a "science teacher retention index" which is a ratio of retained secondary science teachers to the district student population. These are preliminary analyses, yet they demonstrate the power and potential these data sets offer a fresh approach to questions of teacher retention.

Currently, there are still some limitations on the use of these data. Not all states collect exactly the same information, and longitudinal approaches to analysis often entail aligning the changes in year-to-year data categories (for example, race/ethnicity was category in many states prior to 2010, but multiple categories after. Issues of both data security and the politically sensitive nature of some data (e.g. linking retention to education preparation program) also influence the

availability and public nature of certain data—topics that will be discussed in detail in the final paper. They also depend heavily on correct input at the level of the LEA. Finally, some existing data systems, such as those in Texas and North Carolina, de-identify individual teachers in all public facing data, which could impact certain lines of inquiry.
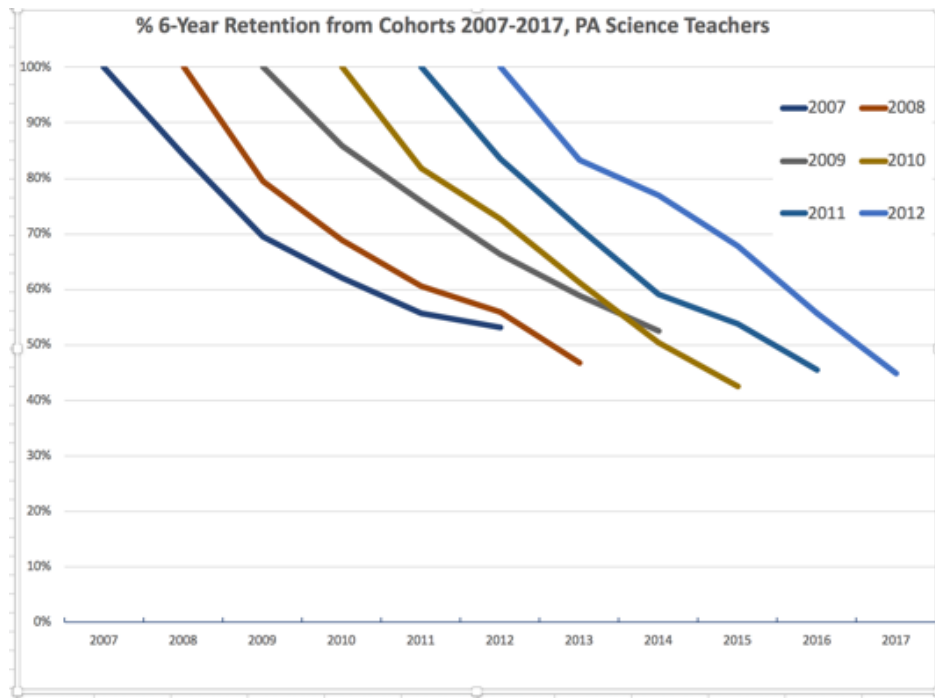

Figure 1. Retention rates for cohorts of new PA science teachers from 2007-2017

AUN	LEA Name	Student population	Index	Number of Science Teachers retained 5+ years 2007-2017
116191757	Columbia-Montour AVTS	645	9.30	3
108112003	Ferndale Area SD	691	7.24	3
122097604	New Hope-Solebury SD	1432	5.59	4
104435303	Reynolds SD	1105	5.43	3
116495103	Mount Carmel Area SD	1497	5.34	4
118403903	Lake-Lehman SD	1789	4.47	4
119355503	Mid Valley SD	1814	4.41	4
122091457	Bucks County Technical High Scho	1396	4.30	3
101633903	McGuffey SD	1576	3.81	3
111312503	Huntingdon Area SD	1860	3.76	4
112679107	York Co School of Technology	1679	3.57	3
124159002	West Chester Area SD	11923	3.35	20
112676503	Southern York County SD	3058	3.27	5
107652603	Franklin Regional SD	3443	3.19	6
107657503	Southmoreland SD	1889	3.18	3
119351303	Carbondale Area SD	1588	3.15	3
103026303	Montour SD	2870	3.14	5
103029902	Woodland Hills SD	3529	3.12	6
114069353	Wyomissing Area SD	1936	3.10	3
124150004	Pennsylvania Leadership Charter S	2594	3.08	4
106616203	Oil City Area SD	1954	3.07	3
129545003	North Schuylkill SD	1958	3.06	3
107651603	Derry Area SD	1971	3.04	3
126513452	Olney Charter High School	1984	3.02	3
121136503	Palmerton Area SD	1715	2.92	3
101301403	Central Greene SD	1730	2.89	3
124152003	Downingtown Area SD	12794	2.74	18

Figure 2. LEA Secondary Science Teacher Retention Index from PA 2007-17

Scholarly significance of the study

We wish to conclude this paper by offering a list of potential avenues of research about teacher retention that could be informed by using state-level school staffing data. Such research includes examining factors related to teacher retention such as the presence of additional certifications (e.g. Teaching students with disabilities, subject area, bilingual/bicultural certifications), salary, demographic match between teacher and the LEA, and the effect of a leave of absence on retention rates. These data could also be used to examine equity issues, such as gendered salary gaps resulting from mobility (we found many cases where teachers

10

moving to new schools had incorrect years of experience listed), or whether second career teachers, defined perhaps as those who begin teaching after a particular age, are retained at comparable rates to their younger peers. These findings would have implications for recruitment, preparation, induction and mentoring and would be valuable for teacher educators, LEA administrators, and policymakers.

Furthermore, this paper connects strongly to the AERA 2020 program theme because of the connection between state-level stakeholders who shape the data collected about teachers and researchers who can pose new questions of these data that help inform future practice and policy toward strengthening teacher retention efforts.

References

Achinstein, B., & Barrett, A. (2004). (re)framing classroom contexts: How new teachers and mentors view diverse learners and challenges of practice. *Teachers College Record, v106*(n4), p716.

Achinstein, B., Ogawa, R. T., Sexton, D., & Freitas, C. (2010). Retaining teachers of color: A pressing problem and a potential strategy for "hard-to-staff" schools. *Review of Educational Research, 80*(1), 71-107. doi: 10.3102/0034654309355994

Bang, E., & Luft, J. (2014). Exploring the written dialogues of two first-year secondary science teachers in an online mentoring program. *Journal of Science Teacher Education, 25*(1), 25-51. doi: 10.1007/s10972-013-9362-z

Bianchini, J. A., & Brenner, M. E. (2010). The role of induction in learning to teach toward equity: A study of beginning science and mathematics teachers. *Science Education, 94*(1), 164-195.

Borman, G. D., & Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of educational research, 78*(3), 367-409.

Boser, U. (2012). Race to the top: What have we learned from the states so far?: A state-by-state evaluation of race to the top performance Washington, DC: Center for American Progress. Retrieved from https://cdn.americanprogress.org/wp-content/uploads/issues/2012/03/pdf/rtt_states.pdf

Boyd, D. J., Grossman, P., Lankford, H., Loeb, S., Michelli, N. M., & Wyckoff, J. (2006). Complex by design: Investigating pathways into teaching in New York city schools. *Journal of Teacher Education, 57*(2), 155-166. doi: 10.1177/0022487105285943

Ceven McNally, J. (2016). Learning from one's own teaching: New science teachers analyzing their practice through classroom observation cycles. *Journal of Research in Science Teaching, 53*(3), 473-501. doi: 10.1002/tea.21253

Dawson, P. (2014). Beyond a definition: Toward a framework for designing and specifying mentoring models. *Educational Researcher, 43*(3), 137-145. doi: 10.3102/0013189X14528751

Flores, S. M., Park, T. J., Viano, S. L., & Coca, V. M. (2017). State policy and the educational outcomes of English learner and immigrant students: Three

administrative data stories. *American Behavioral Scientist, 61*(14), 1824-1844. doi: 10.1177/0002764217744836

Ingersoll, R. M. (1997). Teacher turnover and teacher quality: The recurring myth of teacher shortages, Editorial, *Teachers College Record,* p. 41. Retrieved from https://www.tcrecord.org/content.asp?contentid=10602

Ingersoll, R. M. (2007). The science and mathematics teacher shortage: Fact and myth. *NSTA Reports!, 18*(9), 6-7.

Ingersoll, R. M. (2011). Do we produce enough mathematics and science teachers? *Phi Delta Kappan, 92*(6), 37-41. doi: 10.1177/003172171109200608

Ingersoll, R. M., & May, H. (2011). The minority teacher shortage: Fact or fable? *Phi Delta Kappan, 93*(1), 62-65.

Ingersoll, R. M., & May, H. (2012). The magnitude, destinations, and determinants of mathematics and science teacher turnover. *Educational Evaluation and Policy Analysis, 34*(4), 435-464.

Ingersoll, R. M., Merrill, L., & May, H. (2016). Do accountability policies push teachers out? Sanctions exacerbate the teacher turnover problem in low-performing schools--but giving teachers more classroom autonomy can help stem the flood. *Educational Leadership, 73*(8), 44-49.

Ingersoll, R. M., & Smith, T. M. (2003). The wrong solution to the teacher shortage. *Educational Leadership, 60*(8), 30.

Lee, E. (2006). Making equity explicit: A professional development model for new mentors. In B. Achinstein & S. Z. Athanases (Eds.), *Mentors in the making: Developing new leaders for new teachers* (pp. 55-65). New York: Teachers College Press.

Luft, J. A., Firestone, J. B., Wong, S. S., Ortega, I., Adams, K., & Bang, E. (2011). Beginning secondary science teacher induction: A two-year mixed methods study. *Journal of Research in Science Teaching, 48*(10), 1199-1224. doi: 10.1002/tea.20444

Mandel, Z. R., Fuller, E., & Pendola, A. (2018). *Production, placement, and retention of secondary STEM teachers of color: A case study of texas.* Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

Marinell, W. H., & Coca, V. M. (2013). Who stays and who leaves? Findings from a three part study of teacher turnover in nyc middle schools. New York: from

https://steinhardt.nyu.edu/scmsAdmin/media/users/sg158/PDFs/ttp_synthesis/TTPSynthesis_Report_March2013.pdf

McFarland, J., Hussar, B., Zhang, J., Wang, X., Wang, K., Hein, S., Diliberti, M., Forrest Cataldi, E., Bullock Mann, F., and Barmer, A. (2019). *The condition of education 2019*. (NCES 2019-144). Washington, DC: Retrieved from https://nces.ed.gov/pubs2019/2019144.pdf.

McGinnis, J. R., Parker, C., & Graeber, A. O. (2004). A cultural perspective of the induction of five reform-minded beginning mathematics and science teachers. *Journal of Research in Science Teaching, 41*(7), 720-747. doi: 10.1002/tea.20022

Natriello, G., & Zumwalt, K. (2017). Pathways to a profession. *Teachers College Record, 119*(14), 145-147.

Roehrig, G. H., & Luft, J. A. (2006). Does one size fit all? The induction experience of beginning science teachers from different teacher-preparation programs. *Journal of Research in Science Teaching, 43*(9), 963-985.

Saka, Y., Southerland, S., Kittleson, J., & Hutner, T. (2013). Understanding the induction of a science teacher: The interaction of identity and context. *Research in Science Education, 43*(3), 1221-1244. doi: 10.1007/s11165-012-9310-5

Simon, N. S., & Johnson, S. M. (2015). Teacher turnover in high-poverty schools: What we know and can do. *Teachers College Record, 117*(3), 1-36.

Spellings, M. (2007). *Final guidance on maintaining, collecting, and reporting racial and ethnic data to the U.S. Department of Education* (72 Fed. Reg. 59267). Federal Register: Retrieved from https://title2.ed.gov/Public/TA/Guidance.pdf.

Zeichner, K. (2006). Studying teacher education programs: Enriching and enlarging the inquiry. In C. Conrad & R. C. Serlin (Eds.), *The sage handbook for research in education: Engaging ideas and enriching inquiry* (pp. 79-93). Thousand Oaks: Sage Publications.

Zumwalt, K., Natriello, G., Randi, J., Rutter, A. L., & Sawyer, R. (2017). Recruitment, preparation, placement, and retention of alternate route and college-prepared teachers: An early study of a New Jersey initiative. *Teachers College Record, 119*(14), 28-40.

Zumwalt, K., Randi, J., Rutter, A. L., & Sawyer, R. (2017). First 11 years: 25 exemplar alternate route and college-prepared k-12 teachers. *Teachers College Record, 119*(14), 42-55.