Science AAAS

The Effects of Experience and Attrition for Novice High-School Science and Mathematics Teachers

Gary T. Henry *et al.*

Science **335**, 1118 (2012);
DOI: 10.1126/science.1215343

distribution continuous, we computed an "achievement rank" by combining the number of the stage reached and the number of successful manipulations (see supporting online material).
28. G. Csibra, G. Gergely, in *Processes of Change in Brain and Cognitive Development*, Y. Munakata, M. Johnson, Eds. (Oxford Univ. Press, Oxford, 2006), pp. 249–274.
29. D. Wood, J. S. Bruner, G. Ross, *J. Child Psychol. Psychiatry* **17**, 89 (1976).
30. R. L. Kendal, J. R. Kendal, W. Hoppitt, K. N. Laland, *PLoS ONE* **4**, e6541 (2009).
31. A. Whiten *et al*., *Curr. Biol.* **17**, 1038 (2007).
32. M. Dindo, B. Thierry, A. Whiten, *Proc. R. Soc. B* **275**, 187 (2008).
33. M. Tomasello, M. Carpenter, J. Call, T. Behne, H. Moll, *Behav. Brain Sci.* **28**, 675 (2005).

Acknowledgments: Supported by the CULTAPTATION project (European Commission contract FP6-2004-NESTPATH-043434), a European Research Council Advanced Grant (EVOCULTURE, 232823) (K.N.L.), and a Royal Society Dorothy Hodgkin Fellowship (R.L.K.). Chimpanzees at M. D. Anderson Cancer Center were supported by NIH Cooperative Agreement RR-15090. We thank S. Pavonetti, A. Whiten, C. Caldwell, C. Tennie, M. Tomasello, C. van Schaik, the CCBC/EARG journal club (Durham), and members of the Laland lab for advice and/or useful comments on earlier drafts of our manuscript, and J. Zampol for drawing Fig. 1. All research was approved by the ethics committee at the University of St. Andrews and the relevant committees at the research facilities and follows the relevant legislation in the UK, USA, and France. Data are available from the authors upon request.

Supporting Online Material
www.sciencemag.org/cgi/content/full/335/6072/1114/DC1
Materials and Methods
Figs. S1 and S2
Tables S1 to S6

14 September 2011; accepted 23 January 2012
10.1126/science.1213969

The Effects of Experience and Attrition for Novice High-School Science and Mathematics Teachers

Gary T. Henry,[1]* C. Kevin Fortner,[2] Kevin C. Bastian[3]

Because of the current high proportion of novice high-school teachers, many students' mastery of science and mathematics depends on the effectiveness of early-career teachers. In this study, which used value-added models to analyze high-school teachers' effectiveness in raising test scores on 1.05 million end-of-course exams, we found that the effectiveness of high-school science and mathematics teachers increased substantially with experience but exhibited diminishing rates of return by their fourth year; that teachers of algebra 1, algebra 2, biology, and physical science who continued to teach for at least 5 years were more effective as novice teachers than those who left the profession earlier; and that novice teachers of physics, chemistry, physical science, geometry, and biology exhibited steeper growth in effectiveness than did novice non–science, technology, engineering, and mathematics teachers.

In the past two decades, the teacher labor market has dramatically changed in response to more employment opportunities for women, increased demand for teachers, and policies opening new pathways into the profession (*1*). For instance, the modal value of experience for U.S. teachers dropped from 15 years in 1987–1988 to 1 year in 2007–2008 (*2*). Additionally, turnover for beginning teachers is high: After just 5 years, nearly 50% of all novice teachers have exited the profession (*3*). This churn of beginning teachers in and out of public schools results in more students, particularly poor and/or ethnic minority students, being taught by novice teachers (*4*, *5*), and that in turn leads to reduced student achievement (*6*–*11*). Investigating the consequences of these teacher labor market conditions for high-school students' science and mathematics achievement—specifically, the effects of experience and attrition among novice teachers—will shed light on challenges facing education and career preparation in science and mathematics.

We quantified the growth in effectiveness of high-school science and mathematics teachers and the effects of those teachers who exit public school classrooms. We analyzed effectiveness using scores on standardized tests given to high-school students in three mathematics courses and four science courses. We define teachers' effectiveness in terms of the increases in their students' test scores, adjusted for the prior achievement of the individual students and for other student, classroom, and school covariates. Prior research shows that the average effectiveness of novice teachers increases during their first 3 years and flattens thereafter (*12*–*14*); and that after differences in effectiveness that are attributable to experience are removed, less effective teachers are more likely to exit the profession (*10*, *15*, *16*). We extended this research to investigate teacher effectiveness in specific high-school courses and addressed three questions: (i) To what extent do novice high-school science and mathematics teachers become more effective with additional experience? (ii) Are novice high-school science and mathematics teachers who exit public schools more or less effective than those who stay? [There is little published research on where teachers who exit public schools are subsequently employed, but it suggests that leaving for higher-paying jobs in the private sector is relatively rare (*17*).] (iii) Do the rates of change in effectiveness for high-school science and mathematics teachers vary by course?

Student outcomes are related to variations in many school-related factors, including leadership; an orderly environment; high student expectations; a focus on student outcomes; a positive school culture; parental involvement; and, most closely, to teachers' effectiveness (*18*, *19*). Teachers are the most important school-related variable explaining variation in student achievement (*12*, *18*, *20*), and teacher experience positively affects student performance (*6*–*11*). Most gains in effectiveness occur in the first 3 years of teachers' careers, with minimal increases thereafter (*13*).

It is likely that through teaching experience, trial and error, professional development, mentoring, and/or collaboration with fellow educators, teachers learn rapidly during their first few years on the job. However, some of the average increases in effectiveness that have been attributed to experience may be a statistical artifact caused by the exit of less effective early-career teachers, thereby overstating the year-to-year differences in the statistics related to experience. Recent research supports this second explanation, finding that exiting teachers are less effective than comparable teachers who remain in the profession (*10*, *15*, *16*). Here we disentangle the effects of teacher development from differential attrition among high-school teachers of science and mathematics courses.

To assess the effectiveness and attrition of novice science and mathematics teachers, we developed a data set from North Carolina containing end-of-course test scores for seven science and mathematics courses—algebra 1, algebra 2, geometry, biology, chemistry, physical science, and physics—and three other courses—English 1, U.S. history, and civics/economics—which we grouped together for the purposes of our analysis and label non-STEM (science, technology, engineering, and mathematics) courses. Our study sample included all teachers in tested subjects with less than 5 years of experience employed in any regular North Carolina public high school from 2005–2006 through 2009–2010. The most crucial feature of the data set is that students and teachers were linked on the basis of actual classroom rosters, which allowed us to match approx-

[1]Department of Public Policy and Education Policy Initiative at Carolina, University of North Carolina at Chapel Hill, Abernethy Hall, Campus Box 3435, Chapel Hill, NC 27599–3435, USA. [2]Department of Educational Policy Studies, College of Education, Georgia State University, 30 Pryor Street, Atlanta, GA 30303–3083, USA. [3]Department of Public Policy and Education Policy Initiative at Carolina, University of North Carolina at Chapel Hill, Abernethy Hall, Campus Box 3279, Chapel Hill, NC 27599–3435, USA.

*To whom correspondence should be adddressed. E-mail: gthenry@unc.edu

imately 93% of high-school teachers to students in their courses; to construct student, teacher, and classroom covariates; and to account for students who had multiple teachers for any course. In total, the sample included 1.05 million test scores, 624,842 unique students, and 7961 unique teachers with less than 5 years of experience.

The outcome variable analyzed was the students' score on a standardized exam: the North Carolina End-of-Course Tests. These exams are designed to test students' knowledge and skills based on the North Carolina Standard Course of Study objectives for each course, which follow the frameworks of the National Assessment of Educational Progress and recommendations from national standard-setting groups (*21*, *22*). The items on these tests covered a range of cognitive skills, from remembering to analyzing and evaluating; included four potential response options; and were designed to cover a range of difficulty from easy (25%) to medium (50%) to hard (25%). Multiple versions of each test with different items reduce teachers' ability to predict the items, while providing equivalent scores across different versions. For example, there were 12 versions of the chemistry test, each of which included 92 unique items, meaning that a total of 1104 items were in use during the study period (*21*). To remove any year-to-year differences in test scores and to estimate effects in standard deviation units (SDUs), we standardized all tests within course and year by centering observations on the mean and dividing by the standard deviation.

To assess the development and attrition of novice science and mathematics teachers, we estimated two sets of subject-specific models, in which focal development variables were four dichotomous experience indicators coded 1 for teachers with the designated experience level (1, 2, 3, or 4 years, respectively) and 0 for teachers with other experience levels (comparisons were to teachers with 0 prior years of experience). The two sets of models included different focal variables to measure the impact of teacher attrition and separate its effect from teacher development. Based on certified salary files supplied by the North Carolina Department of Public Instruction, the focal variable in model 1, labeled "leaves within 5 years," was coded 1 for teachers who leave North Carolina public schools before beginning a sixth year of service and 0 for all others; the focal variable in model 2, labeled "last year," was coded 1 for teachers in a given year who will not be paid as teachers in the following school year and 0 for all others. In model 1, the coefficient on "leaves within 5 years" provided an estimate of the overall average effectiveness of exiting teachers, while netting out the effects of attrition [see eq. 1 in the supporting online material (SOM)]. In model 2, the coefficient on the "last year" indicator estimated the average effectiveness of novice teachers in their final year, while removing the effect of teachers who will leave at the end of the year (which is commonly known as the marginal rate of return to experience). In addition, model 2 included interaction terms between experience and "last year" to allow the average effectiveness of departing teachers to vary based on the year of exit (see eq. 2 in the SOM).

To estimate differences in the rates of change of teachers' effectiveness by course, we combined all courses in a third model and substituted a continuous measure of experience and its squared term into the value-added models. In model 3, we interacted both experience measures with course indicator variables to compare slopes and rates of change by subject (see eq. 3 in the SOM). Results from this model compare the average returns to experience for novice science and mathematics teachers to those of novice teachers in non-STEM courses. All model specifications use value-added models with student, classroom, and school covariates, including students' prior test scores (Table 1).

We find (Table 2) that teachers of all four science subjects experience gains of at least 0.11 SDU ("2nd Year Teachers") between their first and second year of teaching, as compared to lesser gains of 0.06 to 0.09 SDU for all three mathematics courses and non-STEM courses. Second-year physics teachers post the largest gains of 0.38, with chemistry, physical science, and biology teachers posting gains of 0.17, 0.16, and 0.12 SDU, respectively. For all teachers, returns to experience diminish rapidly. The effectiveness of novice algebra 2, geometry, biology, and chemistry teachers peaks in their fourth year, whereas algebra 1, physics, and non-STEM teachers continue to increase in effectiveness through their fifth and final year that is included in this data set. In the middle panels of Table 2, the coefficients on the experience variables quantify the marginal rates of return to experience, separating out the effects of teachers who leave the next year, which are either larger or smaller than the coefficients on experience in model 1, depending on the effectiveness of teachers who exited.

Comparing average effectiveness, novice teachers who will remain teaching beyond 5 years are more effective in four of seven science and mathematics courses than teachers who will leave before completing 5 years (Table 2). In courses where differences are significant, the average difference in effectiveness for teachers who exit within 5 years ranges from –0.102 for physical science

Table 1. Covariates used in analyses.

Student Covariates	Classroom Covariates	School Covariates
1) Prior test scores (mathematics and reading)	1) Class size	1) School size
2) Classmates' prior test scores	2) Heterogeneity of prior achievement within the class	2) School size squared
3) Days absent	3) Advanced curriculum	3) Violent acts per 1,000 students
4) Structural mobility	4) Remedial curriculum	4) Total per-pupil expenditures
5) Within year mobility	5) Teacher out-of-field status	5) Average district teacher supplement
6) Other between year mobility		6) Racial/ethnic composition
7) Race/ethnicity		7) Concentration of poverty
8) Gender		
9) Poverty status		
10) Gifted		
11) Disability		
12) Currently limited English proficient		
13) Was limited English proficient		
14) Overage for grade		
15) Underage for grade		
16) Model 3 only: High school EOC exam (non-STEM courses as the reference group)		

Table 2. Average effectiveness of novice teachers who stay and who exit by course (focal variables only). All results are in comparison to first-year teachers who stay in North Carolina public schools for at least a second year. *, **, and *** indicate significance at the $P < 0.05$, $P < 0.01$, and $P < 0.001$ levels, respectively.

Focal Variables	Algebra 1	Algebra 2	Geometry	Biology	Chemistry	Physical Science	Physics	Non-STEM
2nd Year Teachers	0.056***	0.058*	0.091***	0.115***	0.168**	0.158***	0.378***	0.058***
3rd Year Teachers	0.066***	0.082**	0.125***	0.114***	0.197***	0.159***	0.311***	0.081***
4th Year Teachers	0.067***	0.098***	0.160***	0.122***	0.242***	0.156***	0.321***	0.084***
5th Year Teachers	0.084***	0.091**	0.142***	0.105***	0.236***	0.114**	0.393***	0.089***
Leaves within 5 Years	-0.071***	-0.074**	0.018	-0.036 *	-0.013	-0.102**	0.068	-0.013
2nd Year Teachers	0.050***	0.045	0.076***	0.104***	0.129*	0.101**	0.276**	0.051***
3rd Year Teachers	0.061***	0.083**	0.129***	0.106***	0.159**	0.145***	0.343***	0.079***
4th Year Teachers	0.075***	0.104***	0.133***	0.128***	0.233***	0.138**	0.306***	0.085***
5th Year Teachers	0.092***	0.102**	0.137***	0.110***	0.204***	0.108**	0.370***	0.089***
Last Year	-0.079*	-0.133**	-0.074*	-0.045	-0.192	-0.237***	0.007	-0.052**
2nd Year* Last Year	0.029	0.123	0.125*	0.049	0.248	0.270**	0.270	0.066*
3rd Year* Last Year	0.050	0.016	-0.006	0.048	0.275	0.096	-0.132	0.025
4th Year* Last Year	-0.063	0.018	0.206***	-0.030	-0.131	0.141	0.063	-0.001
5th Year* Last Year	-0.017	-0.017	-0.046	-0.024	0.393	0.041	-0.097	0.010
Cases	163,621	77,999	93,715	125,643	30,452	58,863	4,878	494,970
1st Year Teachers	882	316	364	455	146	293	53	1,479
2nd Year Teachers	848	324	378	507	199	311	60	1,653
3rd Year Teachers	666	268	305	426	155	260	47	1,492
4th Year Teachers	596	272	243	357	137	228	49	1,316
5th Year Teachers	496	208	220	297	128	195	46	1,138
Will Leave	470	184	195	366	116	239	51	904

to –0.036 for biology when compared to teachers with the same level of experience who stay. Exiting geometry, chemistry, and physics teachers are neither more nor less effective than those who stay.

In the middle panels of Table 2, we examine the effectiveness of exiting teachers in their final year of teaching ("last year"). Teachers in four science and mathematics courses—algebra 1, algebra 2, geometry, and physical science—and the non-STEM courses who leave the profession within 5 years are less effective in their last year on the job.

For non-STEM courses, novice teachers exhibit a positive slope and diminishing returns to experience (Table 3). Novice algebra 1 and algebra 2 teachers exhibit patterns similar to those of non-STEM teachers. Novice biology and physical science teachers exhibit higher rates of growth, but also rates of return to experience that dampen more quickly than those of non-STEM teachers. Finally, novice geometry, chemistry, and physics teachers show steeper slopes but diminishing returns to experience similar to those of non-STEM teachers.

Figure 1 shows the average effectiveness of teachers in each of their first 5 years in the profession, by subject, to allow effectiveness trajectories to be directly compared. On average, novice teachers' trajectories plateau or fall by their fifth year. Furthermore, the order of the lines indicates the consequences of the loss of experienced science and mathematics teachers when they are replaced by first-year teachers. These consequences, in terms of student achievement, are most sizable for physics and chemistry, representing losses of approximately 0.25 to 0.4 SDU, on average; then physical science, geometry, and algebra 2, representing losses between 0.1 and 0.2 SDU, on average; and finally, algebra 1, biology, and non-STEM, representing losses of less than 0.1 SDU, on average.

Overall, novice teachers of high-school science and mathematics exhibit significant returns to experience early in their careers that diminish after 4 years of teaching. This indicates that (i)

Table 3. Returns to experience for novice high-school teachers by course. *, **, and *** indicate significance at the $P < 0.05$, $P < 0.01$, and $P < 0.001$ levels, respectively.

Main Effects	Coefficient
Teaching Experience	0.0527***
Teaching Experience Squared	-0.0085***
Interaction Effects	**Coefficient**
Physics × Teaching Experience	0.1820*
Physics × Teaching Experience Squared	-0.0286
Chemistry × Teaching Experience	0.1026*
Chemistry × Teaching Experience Squared	-0.0135
Physical Science × Teaching Experience	0.0860**
Physical Science × Teaching Experience Squared	-0.0184**
Geometry × Teaching Experience	0.0549**
Geometry × Teaching Experience Squared	-0.0083
Biology × Teaching Experience	0.0357*
Biology × Teaching Experience Squared	-0.0094*
Algebra 2 × Teaching Experience	0.0306
Algebra 2 × Teaching Experience Squared	-0.0028
Algebra 1 × Teaching Experience	-0.0060
Algebra 1 × Teaching Experience Squared	0.0023

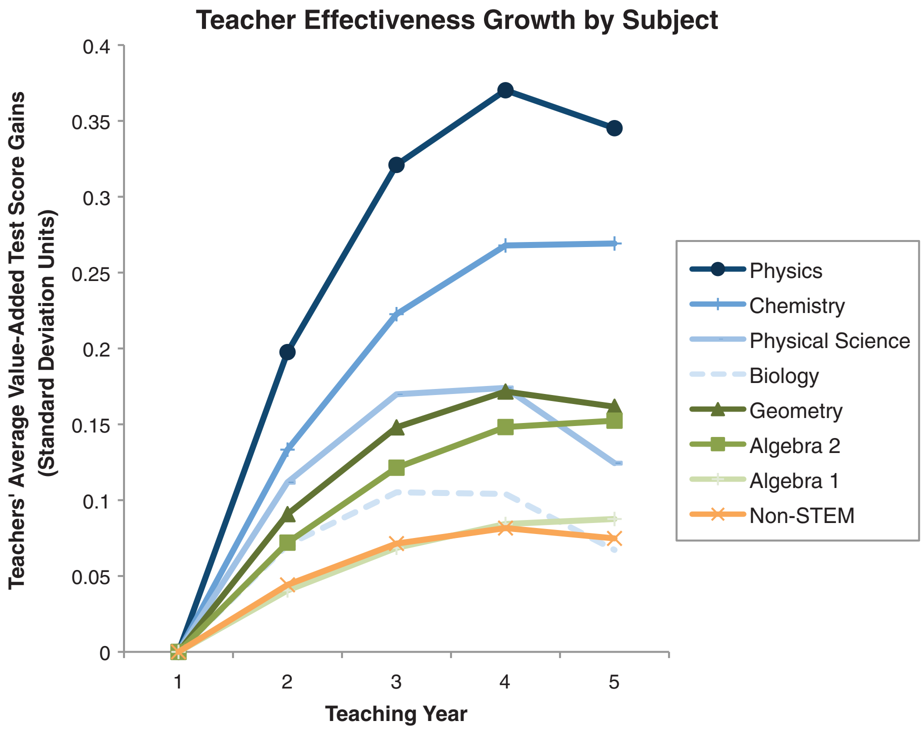

Fig. 1. Average effectiveness of novice high-school teachers.

beginning teachers have a tremendous capacity for improving quickly, and (ii) students of beginning teachers will not achieve at the same levels as students with more experienced teachers. Our results show that increased reliance on novice teachers leads to lower average teacher effectiveness. In light of current teacher labor market conditions, key questions include the following: (i) How to get more effective science and mathematics teachers into high-school classrooms? (ii) Can science and mathematics teachers be better prepared through more preservice experience in classrooms, increased focus on instructional skills, and deeper knowledge of content by the time they begin teaching?

Different high-school subjects show different impacts of teacher turnover. For courses with steeper effectiveness growth curves—physics, chemistry, and geometry—the loss of these experienced teachers has the greatest consequences for student performance. For courses with less steep growth curves—algebra 1, algebra 2, biology, and physical science—the loss of more experienced teachers has less severe consequences. But both cases call for recruiting more able, motivated, and committed teachers. Could these teachers be better screened by evaluating their academic performance, persistence, ability to engage audiences, and projected commitment to teaching, specifically teaching STEM courses to high-school students? Would incentives, such as higher salaries, assignment to fewer courses per year, or paid opportunities for research with university faculty during summers or semester leaves, help retain more of the experienced teachers?

The current churn of the teacher labor market is working against higher student achievement in STEM courses. Although most current education policies that affect teachers do not distinguish between teachers of different types of courses, our results strongly suggest that distinctions between STEM and non-STEM teachers, and even among STEM teachers who teach different courses, may be warranted.

References and Notes

1. M. Bacolod, *Rev. Econ. Stat.* **89**, 737 (2007).
2. R. Ingersoll, L. Merrill, *Educ. Leadership* **67**, 14 (2010).
3. R. Ingersoll, T. Smith, *Educ. Leadership* **60**, 30 (2003).
4. H. Lankford, S. Loeb, J. Wyckoff, *Educ. Eval. Policy Anal.* **24**, 37 (2002).
5. C. Clotfelter, H. Ladd, J. Vigdor, *Econ. Educ. Rev.* **24**, 377 (2005).
6. D. Boyd, P. Grossman, H. Lankford, S. Loeb, J. Wyckoff, *Educ. Finance Policy* **1**, 176 (2006).
7. C. Clotfelter, H. Ladd, J. Vigdor, *Econ. Educ. Rev.* **26**, 673 (2007).
8. C. Clotfelter, H. Ladd, J. Vigdor, *J. Hum. Resour.* **45**, 655 (2010).
9. D. Goldhaber, *J. Hum. Resour.* **42**, 765 (2007).
10. G. T. Henry, K. C. Bastian, C. K. Fortner, *Educ. Res.* **40**, 271 (2011).
11. S. Rivkin, E. Hanushek, J. Kain, *Econometrica* **73**, 417 (2005).
12. J. Rockoff, *Am. Econ. Rev.* **94**, 247 (2004).
13. D. Staiger, J. Rockoff, *J. Econ. Perspect.* **24**, 97 (2010).
14. M. Chingos, P. Peterson, *Econ. Educ. Rev.* **30**, 449 (2011).
15. D. Goldhaber, B. Gross, D. Player, *J. Policy Anal. Manage.* **30**, 57 (2011).
16. J. Krieg, *Econ. Educ. Rev.* **25**, 13 (2006).
17. B. Scafidi, D. Sjoquist, T. Stinebrickner, *Adv. Econ. Anal. Policy* **6**, 1 (2006).
18. C. Teddlie, D. Reynolds, *The International Handbook of School Effectiveness Research* (Falmer Press, London, 2000).
19. D. Reynolds, P. Sammons, B. De Fraine, T. Townsend, J. Van Damme, *Educational Effectiveness Research (EER): A State of the Art Review,* paper presented at the annual meeting of the International Congress for School Effectiveness and Improvement, Cyprus, 2011.
20. J. S. Coleman *et al.*, *Equality of Educational Opportunity* (U.S. Government Printing Office, Washington, DC, 1966).
21. North Carolina Department of Public Instruction, *North Carolina Mathematics Tests. Ed. 3: Technical Report* (Raleigh, NC); downloaded 29 December 2011 from www.dpi.state.nc.us/docs/accountability/reports/mathtechmanualdrafted2.pdf (2008).
22. North Carolina Department of Public Instruction, *North Carolina Science Tests: Technical Report* (Raleigh, NC); downloaded 29 December 2011 from www.dpi.state.nc.us/docs/accountability/testing/reports/eocsciencetechmanual.pdf (2009).

Acknowledgments: We recognize ongoing financial support for this research from the Teacher Quality Research Initiative provided by the University of North Carolina General Administration.

Supporting Online Material

www.sciencemag.org/cgi/content/full/335/6072/1118/DC1
SOM Text
Tables S1 and S2

17 October 2011; accepted 30 January 2012
10.1126/science.1215343