Journal of Science Teacher Education

ISSN: 1046-560X (Print) 1573-1847 (Online) Journal homepage: https://www.tandfonline.com/loi/uste20

National Trends in the Formal Content Preparation of US Science Teachers: Implications of Out-of-Field Teaching for Student Outcomes

Joseph Taylor, Eric Banilower & Grant Clayton

To cite this article: Joseph Taylor, Eric Banilower & Grant Clayton (2020): National Trends in the Formal Content Preparation of US Science Teachers: Implications of Out-of-Field Teaching for Student Outcomes, Journal of Science Teacher Education, DOI: 10.1080/1046560X.2020.1762992

To link to this article: https://doi.org/10.1080/1046560X.2020.1762992

Published online: 21 May 2020.

JOURNAL OF SCIENCE TEACHER EDUCATION
https://doi.org/10.1080/1046560X.2020.1762992

National Trends in the Formal Content Preparation of US Science Teachers: Implications of Out-of-Field Teaching for Student Outcomes

Joseph Taylor [a], Eric Banilower [b*], and Grant Clayton [c*]

[a]Leadership, Research, and Foundations, University of Colorado Colorado Springs, Colorado Springs, Colorado, USA; [b]Horizon Research. Inc., Chapel Hill, North Carolina, USA; [c]Teaching and Learning, University of Colorado Colorado Springs, Colorado Springs, Colorado, USA

ABSTRACT

This study used nationally representative data from the 2011 National Assessment of Educational Progress (8th-grade Science) and the 2018 National Survey of Science and Mathematics Education toward two primary purposes: (a) to examine the association between teachers' formal (university) content preparation in science and student outcomes in science, and (b) to document the prevalence and locality of Out-of-Field (OoF) science teaching in the US. The relationship between teachers' formal science preparation and students' 8th-grade science outcomes was mixed across science disciplines with a statistically significant association being observed for students' earth science outcomes. Teachers' experience teaching science and access to science instructional materials/kits were more strongly associated with student outcomes than was their formal content preparation, with statistically significant associations observed for all student outcomes (physical science, life science, and earth science). The prevalence of OoF science teaching was higher in middle schools than in high schools, as well as more frequently occurring in historically lower achieving and impoverished educational contexts.

KEYWORDS
Out-of-field science teaching; science achievement; instructional materials; national survey of science and mathematics education; NAEP science

Introduction

For decades, Out-of-field (OoF) teaching has posed a significant challenge for all stakeholders of education. As a result, there have been periodic attempts to explore the national prevalence and locality of OoF teaching. During the pre-No Child Left Behind era, the National Center for Education Statistics (NCES) published a report examining the issue (Ingersoll & Gruber, 1996). NCES defined minimally qualified as having at least a college minor in a teaching subject. Using 1990–1991 NCES School and Staffing Survey data and this definition, they explored the prevalence and variation of OoF teaching across settings of different poverty and minority levels. The following is a brief summary of their findings.

OoF teaching was frequent in STEM fields, occurring in as many as 25% of secondary mathematics courses and 56% of physical science courses. Further, schools with higher proportions of students from race/ethnicity groups historically underrepresented in STEM

CONTACT Joseph Taylor jtaylo18@uccs.edu University of Colorado Colorado Springs, 1420 Austin Bluffs Pkwy, Colorado Springs, CO 80918

*Banilower and Clayton are co-equal second authors

(HUS) and levels of poverty were more likely to have OoF teaching in STEM fields. In low-HUS schools, 24.3% of students had OoF teachers in mathematics while high-HUS had 33.6%. Life science showed a similar pattern with the largest group—medium HUS—having 40.2% of students with OoF teachers. Physical science students experienced OoF teaching ranging from 52.7% for medium-HUS schools to 58.6% for high-HUS schools. In high-poverty schools, 32.6% of students were taught mathematics by OoF teachers and 70.6% of physical science students were taught by OoF teachers (p. 16). In 7th grade, 60.4% of life science and 73.8% of physical science students taught by an OoF teacher. They also found a great deal of variation by state.

The National Center for Education Statistics (NCES) publication, the *Condition of Education 2018* (McFarland et al., 2018), documents national patterns around OoF teaching from 2015 that mirror those of the Ingersoll and Gruber report over 20 years ago. In the interim between the two major reports, many educators have assumed that: (a) there is a significant negative relationship between OoF teaching and students' science achievement outcomes, (b) through recent initiatives (e.g., alternative certification programs) the prevalence of OoF science teaching has diminished over time, and (c) OoF science teaching is situated mainly in educational settings that have proportionally more students from race/ethnicity groups historically underrepresented in STEM and/or are economically disadvantaged. However, these assumptions have not been tested with a combination of recent, nationally representative datasets, as well as a high-level of statistical precision.

Purpose and goals

This study extends prior research by rigorously testing current assumptions of OoF science teaching. Specifically, this paper addresses three main research goals:

(1) **Research Goal 1**: Estimate and interpret the strength of the relationship between middle school science teachers' formalized university preparation in science content and student outcomes in science via analysis of the nationally representative *National Assessment of Educational Progress: Science 2011, 8th grade* (National Center for Education Statistics, 2019a)
(2) **Research Goal 2**: Explore the prevalence and locality of OoF teaching in science, using the nationally representative *2018 NSSME+ (*Banilower et al., 2018).
(3) **Research Goal 3**: Based on findings from research goals 1 and 2, draw useful implications/recommendations for science teacher education, workforce development and placement, and equitable learning opportunities for science students.

Research questions

These goals translate into a set of specific research questions that guide the study design and interpretation of its findings:

(1) *Controlling for student demographic characteristics, to what extent is teacher content preparation (formal university education), experience, and access to district-provided*

instructional materials associated with student outcomes for 8th grade science (earth science, life science, and physical science)?

(2) *How prevalent is out-of-field teaching in middle and high school science classes and are there differences in the types of schools and classes in which it occurs?*

Method

Analysis of 2011 NAEP science data (8th grade)

Sample and survey design

The NAEP 2011 8th grade student sample (n = 121,970) was obtained through two-stage random sampling where schools were first randomly sampled, then students randomly sampled from within the randomly sampled schools. Of this student sample, 108,850 students were successfully linked to their respective teachers (n = 12,730) through the teacher-unique testing booklet serial number that is provided in the student data set (all sample sizes rounded to protect respondent anonymity). Sampling weights were applied to this observed student sample to create a nationally representative (estimation) sample of students: population size = 3,761,360. The data used in this study are from the restricted-use data set for NAEP 2011 8th Grade Science.

Student measures

The 2011 8th grade NAEP science outcomes include 5 plausible values (PV) for achievement in each of three science disciplines: earth science, life science, and physical science. Each plausible value was estimated and scaled using Item Response Theory models for each item format (National Center for Education Statistics, 2019b). In addition to achievement variables, demographic and developmental indicator variables such as poverty level, English language learner status (ELL), special education status, and sex were incorporated into the analysis. Eligibility for free or reduced-price lunch (FRL) was used as a proxy for poverty level and presence of an individualized education plan (IEP) was used as the indicator for special education status. Student race indicators were not used as they were strongly correlated with FRL status.

Teacher measures

For the 2011 8th grade administration of NAEP Science, teachers of sampled students responded to a questionnaire about their instructional practices, classroom organization, teaching background, and training. Questions about background and training include items such as teaching experience, degrees, major/minor fields of study, and professional development. Other items seek information about available classroom resources or teachers' control over instructional decisions.

For the purposes of this study we extracted teacher responses to questions regarding their years of experience teaching science and whether they had: an undergraduate major or minor in earth science, life science, or physical science; a graduate major or minor in earth science, life science, or physical science; been issued science instructional materials their school district; been issued science kits by their school district.

For analysis purposes, new binary teacher variables were created using these responses. Specifically, separate binary variables were created indicating whether a teacher had either a graduate major or minor, or an undergraduate major or minor in a given science discipline (MAJMIN). Such a variable was created for earth science, life science, and physical science. Similarly, a single binary variable was created that indicates whether a teacher was issued by their school district either science instructional materials or kits to support instruction (MATKIT).

Statistical analysis

The research questions were addressed using a multiple regression-based approach applied in each of the three disciplines of interest (earth science, life science, and physical science). Specifically, for each discipline, the five plausible achievement values were regressed on student demographic and developmental characteristics, teacher experience, and the two binary indicators for undergraduate/graduate major or minor in the specified discipline and whether the teacher had access to district-provided science instructional materials or kits. Symbolically, using physical science as the example, the regression model for physical science plausible value 1 (PSPV1) was specified as:

$$\mathbf{PSPV1} = \beta_0 + \beta_1(\mathbf{FRL}) + \beta_2(\mathbf{ELL}) + \beta_3(\mathbf{IEP}) + \beta_4(\mathbf{SEX}) + \beta_5(\mathbf{SCIEXP}) + \beta_6(\mathbf{MAJMIN}) + \beta_7(\mathbf{MATKIT}) + \mathbf{e}$$

The analysis was conducted using the plausible values (*pv*) routine within STATA SE v. 14.2 (StataCorp, 2015). The plausible values routine is designed to address key requirements of NAEP analyses: use of multiple plausible value estimates for each outcome, use of final and replicate weights to preserve the national representativeness of the results, and use of variance estimation techniques appropriate for the NAEP survey design (i.e., jackknife estimation).

As the students' plausible values are nested within teachers, ordinary least squares regression will not produce a correct *t* statistic or associated degrees of freedom for accurate statistical significance tests of the teacher-level variables. The plausible values routine is designed to work with the STATA multilevel routine, *xtmixed*, which produces appropriate t-values and degrees of freedom. However, the complexity of the model (5 plausible values, 62 replicate weights, 7 predictors) prohibited the *xtmixed* routine from converging on final estimates.

Thus, an alternative approach was adopted that produces correct estimates. The plausible values routine was conducted in conjunction with the STATA multiple regression command *regress*, which produces the correct regression coefficients for all the variables but overestimates the *t* statistic and associated degrees of freedom for the teacher-level variables.[1] We then applied a correction to the *t* statistic and associated degrees of freedom for each teacher-level variable as described in Hedges (2007). These corrected values result in a Type I error probability (*p*-value) for the teacher-level variables that is appropriately adjusted for autocorrelation of students' scores within teachers. The regression tables in the results section include the corrected *t* statistics, standard errors, and *p* values for teacher-level variables.

[1]Sample STATA syntax for the physical science analysis is provided below to facilitate replication:

pv, pv(PSPV1 PSPV2 PSPV3 PSPV4 PSPV5) jrr jk(2) weight(origwt) rw(srwt*): regress @pv FRL ELL IEP SEX SCIEXP MAJMIN MATKIT [aw = @w]

where origwt is the final sample weight, srwt* ensures use of all 62 replicate sample weights, and jrr jk invokes jackknife estimation.

Limitations

Although 8th grade is an important indicator of science pathways, it includes a number of possible subject-specific—earth science for instance—and general science course offerings. This variation makes defining and quantifying OoF challenging—as who is an OoF or in-field teacher for a general science course? Further it makes quantifying the association of OoF teaching with student outcomes particularly challenging as we have little way of knowing the degree of alignment between the NAEP assessment and coursework of students.

Analysis of 2018 NSSME+ data

Sample and survey design

The 2018 NSSME+ survey—supported by the National Science Foundation and conducted by Horizon Research, Inc. (HRI)—used a stratified, two-stage sample of schools and STEM teachers in the United States. The first stage utilized 2,000 elementary and secondary schools. The second included 10,000 STEM teachers. The Science Teacher Questionnaire measures background, instructional, and school characteristics. Secondary teachers (n = 2,449) were used for this analysis. Further, a single science class was randomly selected for each teacher to respond about. Because the sample involves clustering and unequal probability of selection, HRI utilized WesVar to calculate jackknife and replication weights to compute accurate standard errors for both the overall and secondary samples, The data used in this study are proprietary and not yet publicly released.

Results

In this section we provide our results organized by research question. For 2011 NAEP Science 8th grade, we provide descriptive statistics and regression results separately for each of the three disciplines of interest (earth science, life science, and physical science).

Research question 1: association between teacher characteristics/resources and student outcomes

Descriptive statistics

In Table 1, we provide outcome descriptive statistics to contextualize the measures of association. Specifically, Table 1 includes information on average and extreme data values for the 5 plausible values per discipline, each on a 300-point scale.

Further, Table 2 provides information on the characteristics of the students in the estimation (weighted) sample, as well as characteristics of the teachers linked to those students. Note that although the teachers are linked to a nationally representative sample of students and schools, they are not necessarily a nationally representative sample of teachers. Thus, the proportions related to teacher characteristics in Table 2 should be interpreted caution. With regard to teachers in this study sample, about half have 0–9 years teaching experience, life science is the most prevalent undergraduate or graduate major/minor, and 7 out of 10 have access to district-provided science instructional materials or kits.

Table 1. Descriptive statistics for plausible values in three science disciplines (estimation sample).

	Plausible Value	Minimum	Maximum	Mean	SD
Physical Science	1	0.00	293.16	151.91	33.30
	2	0.00	290.99	151.84	33.41
	3	0.00	284.07	151.88	33.43
	4	0.00	273.45	151.79	33.45
	5	0.00	287.35	151.60	33.32
Earth Science	1	0.00	278.67	151.00	34.04
	2	0.00	278.66	150.96	34.10
	3	0.00	276.42	150.98	34.14
	4	0.00	283.36	150.91	34.15
	5	0.00	276.30	150.76	34.04
Life Science	1	0.00	290.10	152.55	33.85
	2	0.00	278.15	152.49	33.97
	3	0.00	269.71	152.51	33.95
	4	0.00	296.85	152.45	33.97
	5	0.00	285.21	152.30	33.79

Source: NAEP Science 2011 (8th grade)

Table 2. Proportions for categorical or ordinal variables in estimation sample.

					SCIEXP					MAJMIN		
	FRL	ELL	IEP	MALE	0-5 yrs	5-9 yrs	10-19 yrs	20+ yrs		(PS, ES, LS)		MATKIT
%	47%	6%	12%	51%	26%	26%	31%	16%	38%	29%	56%	70%
SE	0.003	0.002	0.001	0.001	0.006	0.007	0.007	0.006	0.007	0.006	0.007	0.006

Source: NAEP Science 2011 (8th grade)

Regression results: association between teacher variables and student outcomes

In Tables 3-5, we provide regression results for each discipline-specific outcome. The pattern of associations for student factors are quite consistent with statistically significant negative associations for the poverty proxy (FRL), ELL status, and special education status, as well as a statistically significant positive association favoring male students. The estimate column of each table provides a measure of association in the raw metric of the plausible value. For instance, the predicted association of FRL with the physical science outcome (Table 3) is that students who are eligible free or reduced-price lunch are predicted to score 21 points lower, on average, on the physical science plausible values, controlling for the other student- and teacher-level factors in the model. For the effect size column, this unstandardized regression coefficient is then standardized on the average standard deviation of the 5 plausible values. In the case of

Table 3. Parameter estimates for the physical science outcome.

	Estimate (β)	SE	*t*	*p*	95% CI Lower	95% CI Upper	Effect Size
Intercept	147.72	0.79	187.32	<.001	-	-	-
FRL	−21.07	0.55	−38.12	<.001	−22.15	−19.99	−0.63
ELL	−36.75	1.52	−24.17	<.001	−39.73	−33.77	−1.10
IEP	−26.43	0.63	−42.03	<.001	−27.66	−25.20	−0.79
SEX	9.54	0.40	23.64	<.001	8.76	10.32	0.29
MAJMIN	0.96	1.32	0.73	.465	−1.63	3.55	0.03
SCIEXP	1.73	0.63	2.73	.006	0.50	2.96	0.05
MATKIT	4.91	1.33	3.68	<.001	2.30	7.52	0.15

Source: NAEP Science 2011 (8th grade)

Table 4. Parameter estimates for the earth science outcome.

	Estimate (β)	SE	*t*	*p*	95% CI Lower	95% CI Upper	Effect Size
Intercept	147.17	0.69	212.15	<.001	-	-	-
FRL	−20.66	0.52	−39.92	<.001	−21.68	−19.64	−0.61
ELL	−36.50	1.44	−25.28	<.001	−39.32	−33.68	−1.07
IEP	−27.23	0.66	−41.39	<.001	−28.52	−25.94	−0.80
SEX	8.43	0.42	20.28	<.001	7.61	9.25	0.25
MAJMIN	2.79	1.27	2.20	.028	0.30	5.28	0.08
SCIEXP	1.61	0.56	2.90	.004	0.51	2.71	0.05
MATKIT	3.98	1.37	2.91	.004	1.29	6.67	0.12

Source: NAEP Science 2011 (8th grade)

Table 5. Parameter estimates for the life science outcome.

	Estimate (β)	SE	*t*	*p*	95% CI Lower	95% CI Upper	Effect Size
Intercept	148.22	0.70	212.22	<.001	-	-	-
FRL	−20.16	0.55	−36.42	<.001	−21.24	−19.08	−0.59
ELL	−37.89	1.48	−25.61	<.001	−40.79	−34.99	−1.12
IEP	−26.83	0.65	−41.25	<.001	−28.10	−25.56	−0.79
SEX	3.62	0.43	8.34	<.001	2.78	4.46	0.11
MAJMIN	0.81	1.11	0.73	.465	−1.37	2.99	0.02
SCIEXP	1.73	0.56	3.09	.002	0.63	2.83	0.05
MATKIT	4.20	1.24	3.39	.001	1.77	6.63	0.12

Source: NAEP Science 2011 (8th grade)

FRL and physical science, the interpretation is similar, only differing in that the predicted disadvantage, on average, to students eligible for lunch aid is 0.63 standard deviations.

The analysis indicates that students of teachers with either a major or minor (MAJMIN) in physical science, earth science, or life science are predicted to score 0.96, 2.79, and 0.81 (respectively) points higher, on average, than will students whose teachers do not have that level of formal content preparation. However, the association is statistically significant only for earth science with a modest effect size of 0.08 standard deviations.

Teaching experience is a significant, positive predictor in all three subjects. For every one-category increase in teacher experience, students are predicted to score 1.73, 1.61, and 1.73, points higher, on average, on the physical science, earth science, and life science outcomes (respectively). The effect size for each subject is also very small—0.05 standard deviations.

Finally, students of teachers with access to district-provided materials or kits are predicted to score 4.91, 3.98, and 4.20 points higher on the physical science, earth science, and life science outcomes (respectively), on average, than students whose teachers do not have that same access. These differences translate into effect sizes of 0.15, 0.12, and 0.12 standard deviations, respectively. We observe that teacher-level factors are somewhat less consistent with the association between teachers' formal content preparation (MAJMIN) and student outcomes being positive for all disciplines but only statistically significant for earth science ($\beta = 2.79$, $p = .028$). Conversely, the association between student outcomes and both teachers' experience teaching science (SCIEXP) and access to district-provided instructional materials/kits (MATKIT) is consistently positive and statistically significant across the physical science ($\beta = 1.73$, $p = .006$; $\beta = 4.91$, $p = < .001$), earth science ($\beta = 1.61$, $p = .004$; $\beta = 3.98$, $p = < .004$), and life sciences ($\beta = 1.73$, $p = .002$; $\beta = 4.20$, $p = .001$).

Research question 2: prevalence of OoF teaching

The 2018 NSSME+ provides nationally representative data about the extent and distribution of out-of-field teaching in science. Because the study did not ask about minors, in-field teaching was defined as having an undergraduate or graduate major in the science discipline of the sampled class. As can be seen in Table 6, out-of-field teaching is very prevalent (perhaps not surprising given that science consists of multiple fields), with almost 7 in 10 high school science teachers and nearly 9 in 10 middle school science teachers teaching at least one class outside their degree field. The percentage of secondary science classes taught by an out-of-field teacher show a similar pattern—88 percent of middle school science classes and 58 percent of high school science classes are taught by a teacher without a degree in the subject (see Table 7). The smaller percentage of high school science classes taught by an out-of-field teacher is most likely due to the fact that life science/biology is the most common degree held by high school science teachers and the most commonly offered high school science class (Banilower et al., 2018). In addition, novice teachers (those in their first 5 years of teaching) are more likely than veteran teachers to be assigned out-of-field classes (see Table 8). Specifically, 66 percent of high school science classes taught by novice teachers are out-of-field compared to 56 percent of classes taught by veteran teachers.

The 2018 NSSME+ data also show differences in out-of-field teaching by area within science and course level. As can be seen in Table 9, 59 percent of middle school life science classes are taught by an out-of-field teacher, compared to 84 percent of physical science and 91 percent of Earth/space science classes. Further, general or integrated science classes

Table 6. Percentage of science teachers with at least one out of field class (se in parentheses).

Grade Level Taught	Out of Field	In Field
Middle School	86 (2.4)	14 (2.4)
High School	69 (1.6)	31 (1.6)

Table 7. Percentage of science class taught by out of field teachers (SE in parentheses).

Grade Level of Class	Out of Field	In Field
Middle School	88 (1.9)	12 (1.9)
High School	58 (1.8)	42 (1.8)

Table 8. Percentage of classes taught by out of field teachers, by teaching experience (SE in parentheses).

	Novice Teachers		Veteran Teachers	
Grade Level of Class	Out of Field	In Field	Out of Field	In Field
Middle School	91 (2.4)	9 (2.4)	87 (2.6)	13 (2.6)
High School	66 (3.3)	34 (3.3)	56 (2.1)	44 (2.1)

Table 9. Middle school science class taught by out of field teachers, by subject area (SE in parentheses).

Subject of Class	Percentage of Classes
Earth/Space Science	91 (2.8)
Life Science/Biology	59 (5.4)
Physical Science	84 (6.9)
General or Integrated Science	100[a]

[a]For this analysis, only those teachers who have degrees in each of the main science disciplines are considered in field for general or integrated science classes. No teachers in the sample met this criterion; thus, a standard error could not be computed.

are quite common in middle schools, addressing all of the major science disciplines. None of these classes are taught by a teacher with expertise in all of those disciplines.

At the high school level, Earth science, environmental science, and physics classes are more likely than chemistry classes, and all of these classes are more likely than biology classes to be taught by an out-of-field teacher (see Table 10). This finding is not surprising given that biology is the most commonly offered high school science class, and thus there is more demand for teachers of biology. Many schools are not able to offer enough sections of subjects like environmental science or physics to have a teacher dedicated to that subject and, for budgetary reasons, often have these classes taught by teachers with a degree in biology or chemistry. The data also show that 72 percent of non-college prep classes are taught by a teacher without a degree in the subject area, compared to 55 percent of 1st year college prep classes and 46 percent of 2nd year (advanced) courses (see Table 11).

The 2018 NSSME+ data also show that the prevalence of out-of-field teaching in high school science varies by a number of other class and school characteristics. As can be seen in Table 12, there are no substantive differences at the middle school level given that the vast majority of classes are taught by teachers without a degree in the subject. At the high

Table 10. High school science class taught by out of field teachers, by subject area (SE in parentheses).

Subject of Class	Percentage of Classes
Earth/Space Science	80 (4.6)
Life Science/Biology	31 (3.0)
Environmental Science/Ecology	90 (4.0)
Chemistry	55 (2.8)
Physics	68 (3.1)
Multi-discipline science courses	100[a]

[a]For this analysis, only those teachers who have degrees in each of the main science disciplines are consider in field for general or integrated science classes. No teachers in the sample met this criterion; thus, a standard error could not be computed.

Table 11. Level of high school science classes taught by out of field teachers (SE in parentheses).

Level of Class	Percentage of Classes
Non-College Prep	72 (2.5)
1st Year College Prep	55 (5.2)
2nd Year College Prep	46 (4.1)

Table 12. Percentage of classes of various types taught by out of field teachers (SE in parentheses).

	Middle School	High School
Prior Achievement Level of Class		
High	88 (3.2)	50 (3.4)
Average/Mixed	89 (2.2)	60 (2.1)
Low	84 (2.2)	68 (6.1)
Percent of Historically Underrepresented Students in Class		
Lowest Quartile	89 (3.2)	55 (2.9)
Second Quartile	87 (4.9)	58 (3.3)
Third Quartile	91 (2.5)	61 (3.5)
Highest Quartile	87 (4.7)	56 (5.0)
Classes in Schools with Various Characteristics		
Community Type		
Rural	85 (4.6)	65 (3.3)
Suburban	91 (1.5)	53 (2.7)
Urban	86 (4.6)	61 (3.0)
Percent of Students in School Eligible for FRL		
Lowest Quartile	91 (1.4)	49 (3.6)
Second Quartile	85 (4.4)	60 (3.1)
Third Quartile	91 (2.7)	60 (4.7)
Highest Quartile	86 (5.5)	65 (3.3)
Region		
Midwest	91 (2.2)	51 (4.6)
Northeast	80 (5.2)	50 (4.6)
South	92 (1.9)	63 (2.5)
West	85 (7.1)	66 (2.8)

school level, classes composed of mostly low-prior-achieving students are much more likely than classes of mostly high-prior-achieving students to be taught by an out-of-field teacher. In terms of school characteristics, high school science classes in rural and urban schools are more likely than those in suburban schools to be taught by out-of-field teachers, and classes in the highest-poverty schools (as measured by percentage of students eligible for FRL) are more likely than those in the wealthiest schools to be taught by out-of -field teachers. Finally, high school science classes are more likely to be taught by an out-of -field teacher in the South and West regions than they are in the Northeast and Midwest. Taken together, these data indicate that the high school students most in need of high-quality teachers are more likely to have a science class taught by a teacher without a degree in the subject of the class.

Discussion

Researchers have documented OoF teaching for over twenty years now (Ingersoll & Gruber, 1996; Lu, Shen, & Poppink, 2007). Despite highly qualified requirements from the No Child Left Behind Act, OoF teaching in the sciences remains widespread. We add nuance to this issue by conducting analyses using nationally representative datasets. NAEP provides nationally representative school and student data and the 2018 NSSME+ provides nationally representative school, class, and teacher data.

We find OoF teaching so pervasive in the middle school that it is the default condition with 88% of classes taught by OoF teachers. In high school, only life sciences have a majority of courses taught by teachers with a subject matter degree. High-poverty and western high schools all had higher percentages of OoF teachers. The distribution of OoF teachers is also not uniform within schools. Non-college prep high school courses have far higher percentages of OoF teachers than advanced courses. In high schools, only half of the classes where students demonstrated high prior achievement had an out OoF teacher while over two-thirds of classes serving low achievers had an OoF teacher. This within-school variation points to essentially a *de facto* tracking mechanism where, because of a lack of qualified teachers, students who need qualified teachers the most are the least likely to be assigned to one.

We find the prevalence of OoF teaching in middle school to be especially alarming. Middle schools have higher percentages of OoF teachers compounded with high percentages of novice teachers teaching OoF. Experience was consistently a significant, positive predictor of achievement net of content knowledge. For Earth science—a frequent middle school subject—having a major or minor was also a significant predictor of achievement yet only 38% teachers in the NAEP sample had a relevant degree. It is difficult to speculate why this association existed for 8th grade Earth science, and no such relationship was detected for life science or physical science. Further study will be needed to test the robustness of these observed differences.

The combined effect of lack of pedagogical and classroom management experience with lack of content knowledge in many classrooms creates a weak foundation for advanced science courses in high school. We caution against experience as a replacement for content knowledge, but having a stronger teacher—as much as experience is a proxy—is clearly better. Perhaps the association of teaching experience with student outcomes is in part due to differences in teacher content knowledge, where the experience-outcomes association is detecting an embedded association between student outcomes and the content knowledge that teachers gain informally through the process of teaching and preparing to teach.

We suspect school administrators are attempting to ameliorate OoF teaching by providing science kits to teachers. Doing so provides a scaffold to the curriculum for novice and OoF teachers and likely aids in creating a consistent experience across classrooms. Access to materials or kits was a consistently significant positive predictor of achievement on 8th grade NAEP science. While clearly helpful, the use of materials/kits should not be seen as a replacement for in-service professional development or formal content preparation through undergraduate or graduate study. Instead, we see the role of high-quality materials/kits as synergistic, providing a mechanism for teacher expertise to manifest for students' benefit. Previous studies have found that the combination of providing teachers with professional development and high-quality instructional materials is associated with higher quality instruction than either component alone (e.g., Banilower, Boyd, Pasley, & Weiss, 2006; Penuel & Gallagher, 2009).

It is important to note also that the stronger associations observed for experience and access to materials, when compared to formal content preparation (i.e., a major or minor in a science discipline), could be specific to the grade level of students and associated sophistication of the science content being taught. Future research might explore whether formalized content preparation becomes more influential on student outcomes, either absolutely or relative to experience or access to materials, when the students are studying and being tested on more advanced science concepts. Twelfth grade NAEP science scores could have provided such a comparison but the 12th-grade NAEP science administration in 2011 did not include a student-linked teacher questionnaire from which teacher experience, formalized content preparation, and access to kits/materials could have been assessed.

Finally, we must call attention to the strong negative association between poverty and student outcomes and the need for teaching strategies for students learning English. Having a high-quality teacher is exceedingly important, but the magnitude of the coefficient for FRL, net of other student and teacher factors, cannot be ignored. In high school, the negative association between poverty and student outcomes is compounded with OoF teaching where the percentage of classes taught OoF in the lowest quartile of FRL was 16% lower than classes in the highest quartile. Similarly, lack of English proficiency remained a consistent predictor of lower achievement on the 2011 8th grade NAEP Science test.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Joseph Taylor http://orcid.org/0000-0002-3753-4888
Eric Banilower http://orcid.org/0000-0002-4224-3786
Grant Clayton http://orcid.org/0000-0002-0423-9663

Funding

This work was supported by the The National Science Foundation [DGE-1642413].

References

Banilower, E. R., Boyd, S. E., Pasley, J. D., & Weiss, I. R. (2006). *Lessons from a decade of mathematics and science reform: A capstone report for the local systemic change through teacher enhancement initiative*. Chapel Hill, NC: Horizon Research, Inc.

Banilower, E. R., Smith, P. S., Malzahn, K. A., Plumley, C. L., Gordon, E. M., & Hayes, M. L. (2018). *Report of the 2018 NSSME+*. Chapel Hill, NC: Horizon Research, Inc.

Hedges, L. V. (2007). Correcting a significance test for clustering. *Journal of Educational and Behavioral Statistics*, *32*(2), 151–179. doi:10.3102/1076998606298040

Ingersoll, R. M., & Gruber, K. (1996). *Out-of-field teaching and educational equality. Statistical Analysis Report*. Washington, DC: US Government Printing Office, Superintendent of Documents, Mail Stop: SSOP, 20402–29328.

Lu, X., Shen, J., & Poppink, S. (2007). Are teachers highly qualified? A national study of secondary public school teachers using SASS 1999–2000. *Leadership and Policy in Schools*, *6*(2), 129–152. doi:10.1080/15700760601168636

McFarland, J., Hussar, B., Wang, X., Zhang, J., Wang, K., Rathbun, A., & Mann, F. B. (2018). *The condition of education 2018* (NCES 2018-144). Washington, DC: National Center for Education Statistics.

National Center for Education Statistics. (2019a). *The nation's report card: Science 2011*. Retrieved from https://nces.ed.gov/nationsreportcard/pubs/main2011/2012465.aspx

National Center for Education Statistics. (2019b). *NAEP technical documentation: Estimation of NAEP score scales*. Retrieved from https://nces.ed.gov/nationsreportcard/tdw/analysis/scaling.aspx

Penuel, W. R., & Gallagher, L. P. (2009). Preparing teachers to design instruction for deep understanding in middle school earth science. *Journal of the Learning Sciences*, *18*(4), 461–508. doi:10.1080/10508400903191904

StataCorp. (2015). *Stata statistical software: Release 14*. College Station, TX: Author.