

advances.sciencemag.org/cgi/content/full/6/25/eaba2091/DC1

Supplementary Materials for

Early-career faculty practice learner-centered teaching up to 9 years after postdoctoral professional development

Nathan C. Emery*, Jessica Middlemis Maher, Diane Ebert-May

*Corresponding author. Email: emeryna1@msu.edu

Published 19 June 2020, *Sci. Adv.* **6**, eaba2091 (2020)
DOI: 10.1126/sciadv.aba2091

The PDF file includes:

Sections S1 to S4
Figs. S1 and S2
Tables S1 and S2
References

Other Supplementary Material for this manuscript includes the following:

(available at advances.sciencemag.org/cgi/content/full/6/25/eaba2091/DC1)

Data S1

Supplemental Materials

Section S1. Confirmatory Factor Analysis output for the Approaches to Teaching Inventory instrument (22 Item, 2-factor model)

> summary(ATI.fit, fit.measures=TRUE, standardized=TRUE)
lavaan 0.6-5 ended normally after 52 iterations

Estimator	ML
Optimization method	NLMINB
Number of free parameters	45
Number of observations	314

Model Test User Model:

Test statistic	559.077
Degrees of freedom	208
P-value (Chi-square)	0.000

Model Test Baseline Model:

Test statistic	2071.682
Degrees of freedom	231
P-value	0.000

User Model versus Baseline Model:

Comparative Fit Index (CFI)	0.809
Tucker-Lewis Index (TLI)	0.788

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-8991.671
Loglikelihood unrestricted model (H1)	-8712.133
Akaike (AIC)	18073.343
Bayesian (BIC)	18242.066
Sample-size adjusted Bayesian (BIC)	18099.339

Root Mean Square Error of Approximation:

RMSEA	0.073
90 Percent confidence interval - lower	0.066
90 Percent confidence interval - upper	0.081
P-value RMSEA <= 0.05	0.000

Standardized Root Mean Square Residual:

SRMR	0.078
------	-------

Parameter Estimates:

Information	Expected
Information saturated (h1) model	Structured
Standard errors	Standard

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
ITTF =~						
A1_1	1.000				0.323	0.441
A1_2	0.077	0.188	0.412	0.681	0.025	0.025
A1_4	1.971	0.284	6.936	0.000	0.636	0.674
A1_6	1.959	0.310	6.322	0.000	0.632	0.548
A2_9	0.410	0.215	1.907	0.056	0.132	0.120
A2_10	1.974	0.303	6.515	0.000	0.637	0.583
A2_11	1.664	0.273	6.086	0.000	0.537	0.509
A2_12	1.418	0.246	5.777	0.000	0.458	0.464
A3_2	2.259	0.317	7.117	0.000	0.729	0.723
A3_5	2.052	0.287	7.151	0.000	0.662	0.733
A3_8	1.385	0.240	5.772	0.000	0.447	0.463
CCSF =~						
A1_3	1.000				0.320	0.403
A1_5	2.118	0.348	6.086	0.000	0.677	0.576
A1_7	1.984	0.329	6.037	0.000	0.635	0.565
A2_8	2.278	0.360	6.322	0.000	0.729	0.636
A2_13	2.288	0.369	6.199	0.000	0.732	0.603
A2_14	0.646	0.188	3.442	0.001	0.207	0.232
A3_1	2.285	0.347	6.586	0.000	0.731	0.719
A3_3	1.701	0.264	6.441	0.000	0.544	0.670
A3_4	1.880	0.309	6.076	0.000	0.602	0.574
A3_6	1.602	0.250	6.397	0.000	0.513	0.657
A3_7	1.796	0.311	5.778	0.000	0.574	0.512

Covariances:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
ITTF ~~						
CCSF	-0.036	0.010	-3.686	0.000	-0.349	-0.349

Variances:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
.A1_1	0.431	0.036	11.957	0.000	0.431	0.805
.A1_2	0.967	0.077	12.528	0.000	0.967	0.999
.A1_4	0.486	0.046	10.533	0.000	0.486	0.546
.A1_6	0.932	0.081	11.510	0.000	0.932	0.700
.A2_9	1.194	0.096	12.495	0.000	1.194	0.986
.A2_10	0.789	0.070	11.304	0.000	0.789	0.660
.A2_11	0.823	0.070	11.698	0.000	0.823	0.741
.A2_12	0.763	0.064	11.878	0.000	0.763	0.785
.A3_2	0.485	0.049	9.896	0.000	0.485	0.477
.A3_5	0.377	0.039	9.736	0.000	0.377	0.462
.A3_8	0.730	0.061	11.881	0.000	0.730	0.785
.A1_3	0.528	0.044	12.129	0.000	0.528	0.838
.A1_5	0.924	0.080	11.499	0.000	0.924	0.668
.A1_7	0.858	0.074	11.557	0.000	0.858	0.681
.A2_8	0.784	0.070	11.121	0.000	0.784	0.596
.A2_13	0.936	0.083	11.342	0.000	0.936	0.636
.A2_14	0.747	0.060	12.412	0.000	0.747	0.946
.A3_1	0.500	0.048	10.302	0.000	0.500	0.483
.A3_3	0.363	0.033	10.831	0.000	0.363	0.550
.A3_4	0.737	0.064	11.511	0.000	0.737	0.671
.A3_6	0.345	0.032	10.948	0.000	0.345	0.568
.A3_7	0.928	0.079	11.793	0.000	0.928	0.738
ITTF	0.104	0.027	3.784	0.000	1.000	1.000
CCSF	0.102	0.030	3.460	0.001	1.000	1.000

> fitMeasures(ATI.fit)

npar	fmin	chisq	df	pvalue	
45.000	0.890	559.077	208.000	0.000	
baseline.chisq	baseline.df	baseline.pvalue	cfi	tli	
2071.682	231.000	0.000	0.809	0.788	
nnfi	rfi	nfi	pnfi	ifi	
0.788	0.700	0.730	0.657	0.812	
rni	logl	unrestricted.logl	aic	bic	
0.809	-8991.671	-8712.133	18073.343	18242.066	
ntotal	bic2	rmsea	rmsea.ci.lower	rmsea.ci.upper	
314.000	18099.339	0.073	0.066	0.081	

rmsea.pvalue	rmr	rmr_nomean	srmr	srmr_bentler
0.000	0.078	0.078	0.078	0.078
srmr_bentler_nomean	crmr	crmr_nomean	srmr_mplus	
srmr_mplus_nomean				
0.078	0.081	0.081	0.078	0.078
cn_05	cn_01	gfi	agfi	pgfi
137.280	146.109	0.856	0.825	0.704
mfi	ecvi			
0.572	2.067			

Discussion of CFA regression coefficients for the ATI instrument

For the “Information Transfer/Teacher-Focused” factor, the lowest standard regression coefficients (SRC) by far were the items “It is important that this course should be completely described in terms of specific objectives that relate to formal assessment items.” (SRC of 0.025) and “I structure my teaching in this course to help students to pass the formal assessment items.” (SRC of 0.12). For the “Conceptual Change/Student-Focused” factor, the lowest SRC was “It is better for students in this course to generate their own notes rather than copy mine.” with an SRC of 0.232. In future studies using the ATI instrument, investigators may want to exercise caution when including these items.

Section S2. Confirmatory Factor Analysis output for the Experiences in Teaching and Learning Questionnaire instrument (7 subscales)

> summary(ETLQ.fit, fit.measures=TRUE, standardized=TRUE)
lavaan 0.6-5 ended normally after 77 iterations

Estimator	ML
Optimization method	NLMINB
Number of free parameters	73
Number of observations	6870

Model Test User Model:

Test statistic	4636.509
Degrees of freedom	278
P-value (Chi-square)	0.000

Model Test Baseline Model:

Test statistic	70250.985
Degrees of freedom	325
P-value	0.000

User Model versus Baseline Model:

Comparative Fit Index (CFI)	0.938
Tucker-Lewis Index (TLI)	0.927

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-243677.537
Loglikelihood unrestricted model (H1)	-241359.282

Akaike (AIC)	487501.073
Bayesian (BIC)	488000.022
Sample-size adjusted Bayesian (BIC)	487768.046

Root Mean Square Error of Approximation:

RMSEA	0.048
90 Percent confidence interval - lower	0.047
90 Percent confidence interval - upper	0.049
P-value RMSEA <= 0.05	0.999

Standardized Root Mean Square Residual:

SRMR	0.046
------	-------

Parameter Estimates:

Information	Expected
Information saturated (h1) model	Structured
Standard errors	Standard

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
DeepApproach =~						
Q4_3	1.000				0.321	0.288
Q4_6	1.604	0.085	18.839	0.000	0.516	0.504
Q5_2	2.165	0.112	19.379	0.000	0.696	0.566
Q5_3	2.210	0.111	19.906	0.000	0.710	0.654
Q5_6	1.647	0.085	19.441	0.000	0.529	0.575
Q6_4	2.012	0.106	19.037	0.000	0.647	0.525
SurfaceApproach =~						
Q4_1	1.000				0.884	0.654
Q4_5	1.137	0.026	43.350	0.000	1.006	0.781
Q6_1	0.435	0.020	21.700	0.000	0.385	0.307
Q6_5	0.924	0.023	39.981	0.000	0.817	0.626

Alignment =~						
ETL1_6	1.000			0.816	0.802	
ETL2_12	1.135	0.014	78.799	0.000	0.927	0.848
ETL4_3	0.865	0.014	63.553	0.000	0.706	0.717
HighQualLearn =~						
ETL2_4	1.000			0.672	0.631	
ETL2_6	1.254	0.025	50.068	0.000	0.842	0.731
ETL2_7	1.278	0.025	51.883	0.000	0.859	0.767
ETL2_10	1.224	0.024	50.240	0.000	0.822	0.734
ETL2_11	1.287	0.025	51.860	0.000	0.865	0.766
OrgStrucContent =~						
ETL1_1	1.000			0.746	0.676	
ETL1_2	1.070	0.020	54.399	0.000	0.798	0.738
ETL1_4	1.246	0.021	58.398	0.000	0.929	0.803
PeerSupport =~						
ETL3_1	1.000			0.795	0.764	
ETL3_4	1.067	0.018	58.762	0.000	0.849	0.795
ETL3_9	1.020	0.017	58.584	0.000	0.811	0.789
Choice =~						
ETL1_3	1.000			1.154	0.835	
ETL1_5	0.775	0.019	40.630	0.000	0.895	0.639

Covariances:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
DeepApproach ~~						
SurfaceApproch	-0.085	0.006	-13.416	0.000	-0.298	-0.298
Alignment	0.095	0.006	15.470	0.000	0.361	0.361
HighQualLearn	0.112	0.007	17.209	0.000	0.520	0.520
OrgStrucContnt	0.088	0.006	15.202	0.000	0.367	0.367
PeerSupport	0.077	0.005	13.990	0.000	0.300	0.300
Choice	0.147	0.009	15.742	0.000	0.396	0.396
SurfaceApproach ~~						
Alignment	-0.351	0.013	-26.510	0.000	-0.487	-0.487
HighQualLearn	-0.260	0.011	-23.554	0.000	-0.437	-0.437
OrgStrucContnt	-0.361	0.013	-27.102	0.000	-0.547	-0.547
PeerSupport	-0.172	0.011	-15.029	0.000	-0.244	-0.244
Choice	-0.222	0.017	-13.063	0.000	-0.218	-0.218
Alignment ~~						
HighQualLearn	0.471	0.013	37.654	0.000	0.859	0.859
OrgStrucContnt	0.581	0.014	40.747	0.000	0.954	0.954
PeerSupport	0.267	0.011	25.372	0.000	0.411	0.411
Choice	0.567	0.017	33.788	0.000	0.601	0.601
HighQualLearn ~~						
OrgStrucContnt	0.401	0.012	34.105	0.000	0.800	0.800

PeerSupport	0.207	0.009	23.213	0.000	0.387	0.387
Choice	0.503	0.015	32.598	0.000	0.649	0.649
OrgStrucContent ~~						
PeerSupport	0.222	0.010	22.377	0.000	0.374	0.374
Choice	0.576	0.017	33.705	0.000	0.669	0.669
PeerSupport ~~						
Choice	0.252	0.015	17.169	0.000	0.275	0.275

Variances:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
.Q4_3	1.140	0.020	56.768	0.000	1.140	0.917
.Q4_6	0.781	0.015	51.629	0.000	0.781	0.746
.Q5_2	1.027	0.021	48.887	0.000	1.027	0.680
.Q5_3	0.674	0.016	43.117	0.000	0.674	0.572
.Q5_6	0.569	0.012	48.440	0.000	0.569	0.670
.Q6_4	1.102	0.022	50.805	0.000	1.102	0.725
.Q4_1	1.049	0.023	44.683	0.000	1.049	0.573
.Q4_5	0.648	0.021	30.504	0.000	0.648	0.390
.Q6_1	1.420	0.025	56.817	0.000	1.420	0.906
.Q6_5	1.035	0.022	46.741	0.000	1.035	0.608
.ETL1_6	0.371	0.008	47.403	0.000	0.371	0.357
.ETL2_12	0.334	0.008	41.866	0.000	0.334	0.280
.ETL4_3	0.471	0.009	52.241	0.000	0.471	0.486
.ETL2_4	0.682	0.013	53.907	0.000	0.682	0.602
.ETL2_6	0.618	0.012	50.392	0.000	0.618	0.466
.ETL2_7	0.517	0.011	48.328	0.000	0.517	0.412
.ETL2_10	0.578	0.012	50.225	0.000	0.578	0.461
.ETL2_11	0.525	0.011	48.360	0.000	0.525	0.413
.ETL1_1	0.662	0.013	52.509	0.000	0.662	0.543
.ETL1_2	0.531	0.011	49.556	0.000	0.531	0.455
.ETL1_4	0.475	0.011	43.808	0.000	0.475	0.355
.ETL3_1	0.452	0.011	40.792	0.000	0.452	0.417
.ETL3_4	0.419	0.011	36.549	0.000	0.419	0.368
.ETL3_9	0.398	0.011	37.397	0.000	0.398	0.377
.ETL1_3	0.578	0.029	19.836	0.000	0.578	0.302
.ETL1_5	1.162	0.026	45.068	0.000	1.162	0.592
DeepApproach	0.103	0.010	10.512	0.000	1.000	1.000
SurfaceApproch	0.782	0.030	26.269	0.000	1.000	1.000
Alignment	0.667	0.017	38.850	0.000	1.000	1.000
HighQualLearn	0.451	0.016	27.830	0.000	1.000	1.000
OrgStrucContnt	0.556	0.018	30.357	0.000	1.000	1.000
PeerSupport	0.632	0.019	33.997	0.000	1.000	1.000
Choice	1.333	0.041	32.163	0.000	1.000	1.000


```

> fitMeasures(ETLQ.fit)
      npar      fmin      chisq      df      pvalue
      73.000      0.337      4636.509      278.000      0.000
baseline.chisq  baseline.df  baseline.pvalue      cfi      tli
      70250.985      325.000      0.000      0.938      0.927
      nnfi      rfi      nfi      pnfi      ifi
      0.927      0.923      0.934      0.799      0.938
      rni      logl  unrestricted.logl      aic      bic
      0.938      -243677.537      -241359.282      487501.073      488000.022
      ntotal      bic2      rmsea  rmsea.ci.lower  rmsea.ci.upper
      6870.000      487768.046      0.048      0.047      0.049
      rmsea.pvalue      rmr      rmr_nomean      srmr      srmr_bentler
      0.999      0.062      0.062      0.046      0.046
srmr_bentler_nomean      crmr      crmr_nomean      srmr_mplus
srmr_mplus_nomean
      0.046      0.048      0.048      0.046      0.046
      cn_05      cn_01      gfi      agfi      pgfi
      472.021      498.526      0.948      0.934      0.751
      mfi      ecvi
      0.728      0.696

```

Discussion of CFA regression coefficients for the ETLQ instrument

For the “Deep Approach” factor, the lowest SRC was the item “I have usually set out to understand for myself the meaning of what we had to learn.” (SRC of 0.288). For the “Surface Approach” factor, the lowest SRC was “I’ve tended to take what we’ve been taught at face value without questioning it much.” with an SRC of 0.307. In future studies using the ETLQ instrument, investigators may want to exercise caution when including these items.

Section S3. Relevant Background Survey Questions. The following is text from the initial background survey that pertain to the results in this paper.

Longitudinal study of early-career faculty teaching practice
 Michigan State University
 University of Wisconsin, Madison

The purpose of this study is to investigate the teaching practices of early-career faculty at diverse institutions nationwide. The proposed study will employ a systems approach to identify the variables at institutional, departmental, and individual levels that significantly impact teaching practice. If you choose to participate, then completing surveys, participating in interviews, and video recording you in your classes over the three years of the study will help us characterize your instructional environment and analyze how these variables might interact with your teaching over time. Results from this study will benefit faculty, departments, and other units interested in supporting teaching reform in biology.

Please be assured that the information from surveys, interviews, videos, and course materials will be kept confidential and you will remain anonymous in any report of research findings. These materials and data will not be used in any way to evaluate you as an individual. No one outside the research team will have access to these items. Your privacy will be protected to the maximum extent allowable by law. Upon request and within these restrictions, results may be made available to you.

Other than possibly needing someone in your classroom to video record you, there are no known risks or discomforts associated with participation in the study. You are free to withdraw from the research at any point, without prejudice or penalty. You will receive \$1000 for participating fully in the 3 year study.

Do you consent to participate in this study?

Yes/No

As we begin our study, we would like to collect some background information about you, focusing on your teaching context and experience. We appreciate your responses to the following questions, which we expect will take you approximately 10-15 minutes to complete.

First, we would like to know more about the course that you plan to use for this study during the 2016-17 academic year. The primary requirement is that the course be designed for undergraduates.

What is the course prefix and number? (e.g., BIO 101)

What is the course title? (e.g., Introductory Biology)

What semester/quarter are you teaching this course?

- Fall 2016
- Winter 2016/17
- Spring 2017
- Summer 2017

What are the start and end dates of your course? This information will help us ensure that we communicate with you at appropriate times for your course schedule.

- Start date (MMDDYYYY)
- End date (MMDDYYYY)

Is this course primarily intended for subject majors (e.g., biology majors) or nonmajors?

- Majors
- Nonmajors
- Both majors and nonmajors

Approximately how many students are enrolled in your course?

How would you describe this course?

- Small lecture-style (<60 students)
- Large lecture-style (>60 students)
- Small lecture with an attached lab component
- Large lecture with an attached lab component
- Small discussion course
- Online course
- Other

If you teach a course with an attached lab component, do you provide the direct instruction for the lab as well?

- Yes, I teach the lab section
- No, graduate and/or undergraduate TA(s) teach the lab section
- No, a different instructor/faculty/lab coordinator teaches the lab section
- Other (please describe)
- My course does not have a lab section

Is this a team-taught course with other faculty? (e.g., you and other faculty teach the same students in the same course either concurrently or consecutively during a single semester)

Yes/No

To what extent do you control your course syllabus?

- I have complete control over the design of my course syllabus, as long as it follows the broad topics indicated in the course description.
- I design my own course syllabus, but there are some external constraints placed on exactly what content is covered (at a finer grain than the broad topics indicated in the course description).
- My course syllabus is largely designed by someone else or in collaboration with other faculty.
- Other (please explain)

To what extent do you design your own class sessions?

- I have complete control over the design of my class sessions.
- I design my own class sessions, but there are some external constraints placed on how content is delivered.
- My class sessions are largely designed by someone else or in concert with other faculty.
- Other (please explain)

To what extent do you design your high-stakes course assessments (e.g., quizzes, exams)?

Choose all that apply.

- I develop the assessments used in the course, including adapting other instructors' assessments to suit my learning outcomes and assessment style.
- I develop the assessments in collaboration with faculty/instructors teaching this or other sections of this course (e.g., a "common exam").

- I use assessments designed/provided by someone else (e.g., questions from the textbook publisher).
- Other (please explain)

The next set of questions collect information about your current role and teaching context.

At what college or university are you employed?

Please describe your current appointment.

- Adjunct faculty
- Visiting scholar
- Tenure-track faculty
- Tenured faculty
- Non-tenure-track faculty/lecturer, with contract renewal yearly
- Non-tenure-track faculty/lecturer, with contract renewal every two or more years
- Other (please describe)

How long have you been in your current position? Please answer in number of years, with a "0" indicating that you are just starting now, and partial years represented with decimals (approximate is fine!).

If you are in a tenurable position, have you already been tenured?

- I am tenured.
- I am not yet tenured, but expect to submit my tenure packet in (please enter year)_____
- I am not in a tenure-track position.

Approximately what percentage (0-100%) of your appointment is dedicated to the following activities?

- Research (e.g., writing, collecting data, mentoring other researchers in your lab group)
- Teaching activities and responsibilities (e.g., course preparation and administration, assisting students with classwork)
- Service (e.g., student advising, committees, outreach)
- Administration (e.g., department chair, field station manager, program administrator)

Total

In what capacities have you previously taught and for how long? Please answer in number of years, using 0 for none and decimals for portions of a year.

- K-12 school teacher
- Graduate teaching assistant
- Postdoctoral scholar
- Faculty/visiting scholar/lecturer/adjunct faculty
- Other (please describe)

Have you participated in any education-related professional development activities within the last 8 years? These could include workshops at your institution, national programs, and education-related professional development offered by your disciplinary society. Please do not include TA training in your response.

Yes/No

What types of professional development have you participated in during the last 8 years? For each, please indicate whether you participated as a graduate student, postdoctoral scholar, and/or faculty/instructor. If you have not participated in that activity during the last 8 years, choose N/A.

Note: please do not include TA training here.

Graduate Student / Postdoc / Faculty/Instructor / N/A

- National teaching professional development programs (e.g. Summer Institutes, FIRST IV)
- Education- related workshops offered at disciplinary society meetings
- Workshops on teaching and learning offered by your institution/department
- Courses on teaching and learning offered by your institution
- Facilitated learning communities or cohort-based training programs in effective teaching offered by your institution
- Informal discussion/reading groups
- Other (please describe)

Choose the highest-engagement professional development activity from above in which you participated. Approximately how many hours of your time were dedicated to that activity? For example, if you have spent most of your teaching professional development time at workshops within your institution, you would count the approximate number of hours spent participating in those workshops during the last 8 years. Note: please do not include time spent in TA training in this total.

Have you developed, led, or managed education research/scholarship in teaching and learning projects?

Yes/No

*Lastly, for the purposes of understanding demographics in this study:
How do you identify your gender? Please write in the box below.*

This concludes our background survey. If there is anything else you think we should know about you or your teaching context and background, please comment in the field below. Otherwise, thank you for providing this valuable information!

Section S4. Discussion of gender results. The following is a description of some of the instructor gender results for the ETLQ and MIST instrument.

Despite past findings that students perceive a classroom differently based on the gender of the instructor (60), we found inconclusive results of instructor gender on ETLQ subscales or MIST scores. The “Surface approach” subscale from the ETLQ was slightly higher for male instructors than female instructors. It is possible that males may be teaching their courses from a surface approach, with more memorizing, unreflective studying, and fragmented knowledge. However, this result is slightly contradicted by the MIST: Course and Self-Reflection subscale, whereby male instructors had higher student-reported values for self-reflection in the classroom. It is unclear if these gender differences for instructors are educationally significant given the contradiction and lack of relationships/differences among the other student survey subscales.

Figure S1. Comparison of RTOP scores for FIRST IV faculty (Dark Blue) and comparison faculty (Light Blue). RTOP scores were significantly different between FIRST IV and comparison faculty (paired Student’s t-test, $t = 5.35$, $P < 0.001$, Cohen’s $d = 0.93$). Each point represents a participant’s RTOP score for a given course. The boxes represent the interquartile range. The categories are as follows: Level 1 (15-30)-straight lecture, no interactions between students and prof; Level 2 (31-45)-lecture with some demonstration and minor student participation; Level 3 (46-60)-significant student engagement with some minds on as well as hands-on involvement; Level 4 (61-75)-active student participation in the critique as well as the carrying out of experiments; Level 5 (76-90)-active student involvement in open-ended inquiry resulting in alternative hypotheses, several explanations, and critical reflection (13)

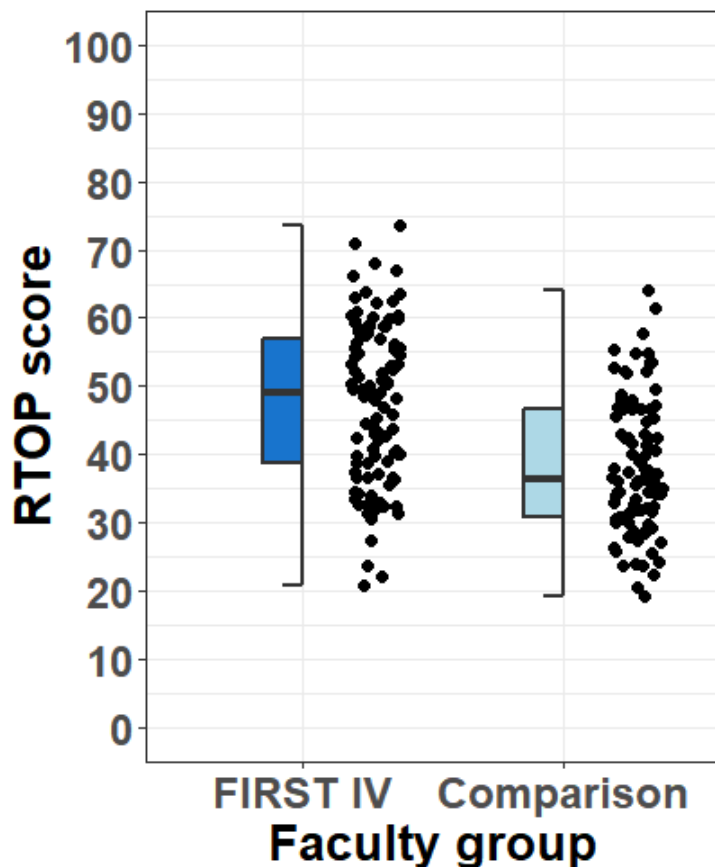


Figure S2. Data collection timeline. A timeline of data collection during the three-year longitudinal study. The Background survey was administered in the beginning of the study. The ATI instrument was sent out in the beginning of a semester/term while the ETLQ and MIST were sent out at the end of a given semester/term. RTOP videos were collected throughout a semester/term.

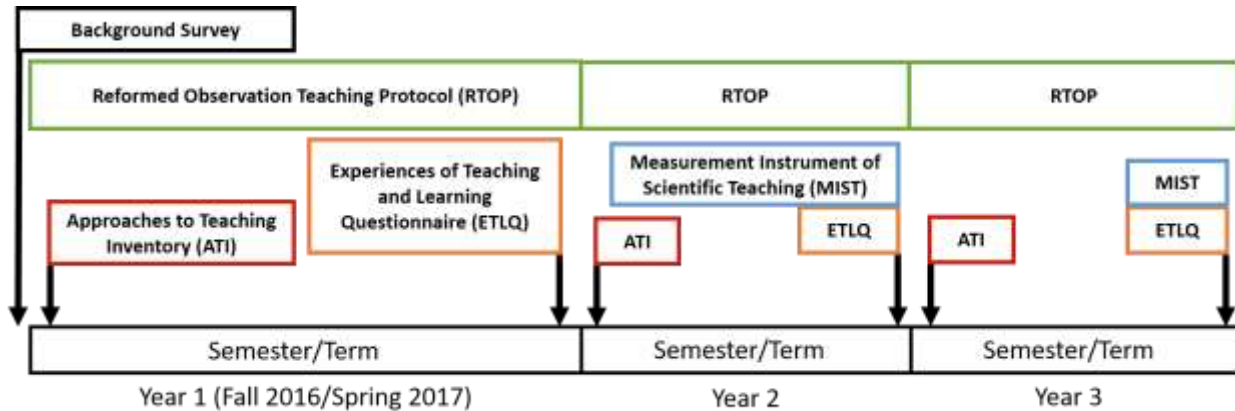


Table S1. Yearly tallies of faculty data collection by measurement instrument. The following is a tally of the number of total faculty participants per year that completed each measurement instrument.

Measurement Instrument	Number of faculty participants (Year 1)	Number of faculty participants (Year 2)	Number of faculty participants (Year 3)
ATI	72	77	62
RTOP	71	70	59
ETLQ	72	76	60
MIST	n/a	76	59

Table S2. Explanation of data columns. The following is an explanation of the data columns for the included data set (Emery_etal_2020_LongitudinalData.csv). Additional data related to this paper may be requested from the authors.

Data Column Title	Explanation
IntID	Institution ID
DeptID	Department ID
PartID	Participant ID
PairID	Pair ID for faculty pairs
FirstID	FIRST IV alumni or colleague
Gender	Self-reported gender identity
Course_YearID	Year of study (1, 2, 3)

Course_UID	Unique ID for a course (.1 = first year of a course, .2 = second year etc)
CarnegieClass	Carnegie classification for an institution
CarnegieUGInstProgram	Carnegie undergraduate program
CarnegieLevel	Carnegie level of an institution (2 or 4 year)
CarnegieControl	Public or private institution
Course_TOY	Time of year of a course
Course_Yr	Year of the course
Student_Type	Majors, Non-majors, or both
Course_Type	Type of course (lecture, lab etc)
Course_TeachType	Who teaches the lab section(s)
Course_Size	Course enrollment
Control_Syllabus	Who has control over the syllabus
Control_Design	Who has control over course design
Control_Assessments	Who has control over assessments
Team_taught	Is this course team-taught
Faculty_Status	Tenure/Non-tenure status
Years_as_faculty	How long have they been a faculty member
Percent_Research	What percent of their appointment is dedicated to research
Percent_Teaching	Percent appointment – teaching
Percent_Service	Percent appointment - service
SumInstructor	Summation of the years as an instructor
SumTeachingExp	Summation of the years of teaching experience
RTOP_FIRST	Average RTOP score from FIRST IV program
CCSF_FIRST	Average CCSF score from FIRST IV program
ITTF_FIRST	Average ITTF score from FIRST IV program
Hist_ETLQ_NResponse	Average number of student responses to the ETLQ from FIRST IV program
Hist_ETLQ_ALS_Deep	Average ETLQ:ALS Deep approach score from FIRST IV program
Hist_ETLQ_ALS_Surface	Average ETLQ:ALS Surface approach score from FIRST IV program
Hist_ETLQ_PTLE_Org_Struc_Content	Average ETLQ:PTLE Organization, Structure, and Content score from FIRST IV program
Hist_ETLQ_PTLE_Alignment	Average ETLQ:PTLE: Alignment score from FIRST IV program
Hist_ETLQ_PTLE_Choice	Average ETLQ:PTLE Choice score from FIRST IV program
Hist_ETLQ_PTLE_Enc_HighQ_Learning	Average ETLQ:PTLE Encouragement of high quality learning score from FIRST IV program
Hist_ETLQ_PTLE_Support_from_students	Average ETLQ:PTLE Support from other students score from FIRST IV program
MIST_AL	MIST: Active Learning
MIST_LG	MIST: Learning Goals

MIST_Inclu	MIST: Inclusivity
MIST_Respons	MIST: Responsiveness to Students
MIST_ED	MIST: Experimental Design and Communication
MIST_DA	MIST: Data Analysis and Interpretation
MIST_CS	MIST: Cognitive Skills
MIST_Reflect	MIST: Course and Self-Reflection
MIST_Score	MIST: Composite score
MIST_NResponses	MIST: Number of student responses
RTOP_Score	RTOP score
N_Responses	Number of student responses to the ELTQ
ALS_Deep	ETLQ:ALS Deep approach score
ALS_Surface	ETLQ:ALS Surface approach score
PTLE_Org_Struc_Content	ETLQ:PTLE Organization, Structure, and Content score
PTLE_Alignment	ETLQ:PTLE: Alignment score
PTLE_Choice	ETLQ:PTLE Choice score
PTLE_Enc_HighQ_Learning	ETLQ:PTLE Encouragement of high quality learning score
PTLE_Support_from_students	ETLQ:PTLE Support from other students
Self_Grade	Average student self-grade (out of 10)
Attendance	Average student-reported attendance
CCSF_Score	ATI: CCSF score
ITTF_Score	AT: ITTF score
Course_ResponseRate	Course student response rate

REFERENCES AND NOTES

1. A. Stes, M. Min-Leliveld, D. Gijbels, P. Van Petegem, The impact of instructional development in higher education: The state-of-the-art of the research. *Educ. Res. Rev.* 5, 25–49 (2009).
2. Y. Steinert, K. Mann, B. Anderson, B. M. Barnett, A. Centeno, L. Naismith, D. Prideaux, J. Spencer, E. Tullo, T. Viggiano, H. Ward, D. Dolmans, A systematic review of faculty development initiatives designed to enhance teaching effectiveness: A 10-year update: BEME Guide No. 40. *Med. Teach.* 38, 769–786 (2016).
3. C. A. Manduca, E. R. Iverson, M. Luxenberg, R. H. Macdonald, D. A. McConnell, D. W. Mogk, B. J. Tewksbury, Improving undergraduate STEM education: The efficacy of discipline-based professional development. *Sci. Adv.* 3, e1600193 (2017).
4. M. R. Connolly, Y. G. Lee, J. N. Savoy, The effects of doctoral teaching development on early-career STEM scholars' college teaching self-efficacy. *CBE Life Sci. Educ.*, 17, ar14 (2018).
5. R. C. Hilborn, The role of scientific societies in STEM faculty workshops: Meeting overview. College Park, MD: American Association of Physics Teachers. (2013).
6. S. E. Brownell, K. D. Tanner, Barriers to faculty pedagogical change: Lack of training, time, incentives, and tensions with professional identity? *CBE Life Sci. Educ.* 11, 339–346 (2012).
7. M. A. Pelch, D. A. McConnell, Challenging instructors to change: A mixed methods investigation on the effects of material development on the pedagogical beliefs of geoscience instructors. *Int. J. STEM Educ.* 3, 5 (2016)
8. O. R. Aragón, S. L. Eddy, M. J. Graham, Faculty beliefs about intelligence are related to the adoption of active-learning practices. *CBE Life Sci. Educ.* 17, ar47 (2018)
9. M. M. Cooper, M. D. Marcos, D. Ebert-May, C. L. Fata-Hartley, S. E. Jardeleza, J. S. Krajcik, J. T. Lavery, R. L. Matz, L. A. Posey, S. M. Underwood, Challenge faculty to transform STEM learning. *Science* 350, 281–282 (2015).
10. D. Ebert-May, T. L. Derting, T. P. Henkel, J. M. Maher, J. L. Momsen, B. Arnold, H. A. Passmore, Breaking the cycle: Future faculty begin teaching with learner-centered strategies after professional development. *CBE Life Sci. Educ.* 14, ar22 (2015).
11. C. Pfund, S. Miller, K. Brenner, P. Bruns, A. Chang, D. Ebert-May, A. P. Fagen, J. Gentile, S. Gossens, I. M. Khan, J. Labov, Summer institute to improve university science teaching. *Science* 324, 470–471 (2009).
12. K. Trigwell, M. Prosser, Development and use of the approaches to teaching inventory. *Educ. Psychol. Rev.* 16, 409–424 (2004).
13. D. Sawada, M. D. Piburn, E. Judson, J. Turley, K. Falconer, R. Benford, I. Bloom, Measuring reform practices in Science and Mathematics classrooms: The reformed teaching observation protocol. *School Sci. Math.* 102, 245–253 (2002).

14. N. Emery, J. M. Maher, D. Ebert-May. Studying professional development as part of the complex ecosystem of STEM higher education. *Innov. High. Educ.* 44, 469–479 (2019).
15. T. L. Derting, D. Ebert-May, T. P. Henkel, J. M. Maher, B. Arnold, H. A. Passmore, Assessing faculty professional development in STEM higher education: Sustainability of outcomes. *Sci. Adv.*, 2, e1501422 (2016).
16. K. N. Pelletreau, J. K. Knight, P. P. Lemons, J. S. McCourt, J. E. Merrill, R. H. Nehm, L. B. Prevost, M. Urban-Lurain, M. K. Smith, A faculty professional development model that improves student learning, encourages active-learning instructional practices, and works for faculty at multiple institutions. *CBE Life Sci. Educ.* 17, es5 (2018)
17. M. T. Owens, G. Trujillo, S. B. Seidel, C. D. Harrison, K. M. Farrar, H. P. Benton, J. R. Blair, K. E. Boyer, J. L. Breckler, L. W. Burrus, D. T. Byrd, Collectively improving our teaching: Attempting biology department-wide professional development in scientific teaching. *CBE Life Sci. Educ.* 17, ar2 (2018)
18. J. Handelsman, D. Ebert-May, R. Beichner, P. Bruns, A. Chang, R. DeHaan, J. Gentile, S. Lauffer, J. Stewart, S. M. Tilghman, W. B. Wood, Scientific teaching. *Science* 304, 521–522 (2004)
19. S. L. Yelon, J. K. Ford, S. Golden, Transfer over time: Stories about transfer years after training. *PIQ* 25, 43–66 (2013).
20. R. T. White, H. J. Arzi, Longitudinal studies: Designs, validity, practicality, and value. *Res. Sci. Educ.* 35, 137–149 (2005).
21. A. Stes, M. Clement, P. Van Petegem, The effectiveness of a faculty training programme: Long-term and institutional impact. *Int. J. Acad. Dev.* 12, 99–109 (2007).
22. A. E. Austin, Promoting evidence-based change in undergraduate science education, in *Fourth committee meeting on status, contributions, and future directions of discipline-based education research* (2011).
23. A. M. Knight, K. A. Cole, D. E. Kern, L. R. Barker, K. Kolodner, S. M. Wright, Long-term follow-up of a longitudinal faculty development program in teaching skills. *J. Gen. Intern. Med.*, 20, 721–725 (2005).
24. M. M. Tennill, M. W. Cohen, 9: Assessing the long-term impact of a professional development program. *To Improve the Academy*, 32, 145–159 (2013).
25. M. Stewart, Making sense of a teaching programme for university academics: Exploring the longer-term effects. *Teach. Teach. Educ.*, 38, 89–98 (2014).
26. G. M. Bodner, Changing how data are collected can change what we learn from discipline-based educational research. *PERC Proc.*, 765, 3–10. (2016).
27. S. R. Porter, M. E. Whitcomb, W. H. Weitzer, Multiple surveys of students and survey fatigue. *New Dir. Inst. Res.*, 121, 63–73 (2004).

28. J. Harshman, M. Stains, A review and evaluation of the internal structure and consistency of the Approaches to Teaching Inventory. *Int. J. Sci. Educ.* 39, 918–936 (2017)
29. N. Entwistle, V. McCune, J. Hounsell, Approaches to study and perceptions of university teaching–learning environments: Concepts, measures and preliminary findings. Edinburgh, Scotland: Enhancing Teaching-Learning Environments in Undergraduate Courses Project, University of Edinburgh, Coventry University, and Durham University (2002).
30. J. T. Lavery, S. M. Underwood, R. L. Matz, L. A. Posey, J. H. Carmel, M. D. Caballero, C. L. Fata-Hartley, D. Ebert-May, S. E. Jardeleza, M. M. Cooper, Characterizing college science assessments: The three-dimensional learning assessment protocol. *PLOS ONE* 11, e0162333 (2016).
31. M. F. Durham, J. K. Knight, B. A. Couch, Measurement Instrument for Scientific Teaching (MIST): A tool to measure the frequencies of research-based teaching practices in undergraduate science courses. *CBE Life Sci. Educ.*, 16, ar67 (2017).
32. D. Ebert-May, T. L. Derting, J. Hodder, J. L. Momsen, T. M. Long, S. E. Jardeleza, What we say is not what we do: Effective evaluation of faculty professional development programs. *Bioscience*, 61, 550–558 (2011).
33. T. Huston, C. L. Weaver, Peer coaching: Professional development for experienced faculty. *Innov. High. Educ.*, 33, 5–20 (2008).
34. C. Henderson, A. Beach, N. Finkelstein, Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature. *J. Res. Sci. Teach.* 48, 952–984 (2011).
35. J. A. Middleton, S. Krause, K. Beeley, E. Judson, J. Ernzen, R. Culbertson, Examining the relationship between faculty teaching practice and interconnectivity in a social network, Paper presented at the 2015 ASEE/IEEE Frontiers in education conference, El Paso, TX (2015).
36. A. E. Austin, Faculty cultures, faculty values. *New Dir. Inst. Res.*, 68, 61–74 (1990).
37. A. E. Austin, Preparing the next generation of faculty: Graduate school as socialization to the academic career. *J. High. Educ.*, 73, 94–122 (2002).
38. K. Viskupic, K. Ryker, R. Teasdale, C. Manduca, E. Iverson, D. Farthing, M. Z. Bruckner, R. McFadden, Classroom observations indicate the positive impacts of discipline-based professional development. *J. STEM Educ. Res.* 2, 201–228 (2019)
39. L. Fung, L. P. Y. Chow, Congruence of student teachers’ pedagogical images and actual classroom practices. *Educ. Res.* 44, 313–321 (2010).
40. T. M. Addy, P. Simmons, G. E. Gardner, J. Albert, A new “class” of undergraduate professors: Examining teaching beliefs and practices of science faculty with education specialties. *J. College Sci. Teach.* 44, 91–99 (2015)

41. I. Ajzen, The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.*, 50, 179–211 (1991).
42. G. Joughin, Assessment, learning and judgement in higher education: A critical review, in *Assessment, Learning and Judgement in Higher Education* (Springer, 2009), pp. 1–15.
43. P. L. Machemer, P. Crawford, Student perceptions of active learning in a large cross-disciplinary classroom. *Active Learn. Higher Educ.* 8, 9–30 (2016).
44. M. F. Durham, J. K. Knight, E. K. Bremers, J. D. DeFreece, A. R. Paine, B. A. Couch, Student, instructor, and observer agreement regarding frequencies of scientific teaching practices using the Measurement Instrument for Scientific Teaching-Observable (MISTO). *Int. J. STEM Educ.*, 5, 31 (2018).
45. C. W. Beck, L. S. Blumer, Alternative realities: Faculty and student perceptions of instructional practices in laboratory courses. *CBE Life Sci. Educ.* 15, ar52 (2016)
46. A. B. Heim, E. A. Holt, Comparing student, instructor, and expert perceptions of learner-centeredness in post-secondary biology classrooms. *PLOS ONE* 13, 10.1371/journal.pone.0200524 (2018)
47. D. A. Greer, A. Cathcart, L. Neale, Helping doctoral students teach: Transitioning to early career academia through cognitive apprenticeship. *Higher Educ. Res. Dev.*, 35, 712–726 (2016).
48. Carnegie Foundation for the Advancement of Teaching, The Carnegie Classification of Institutions of Higher Education, 2010 edition, Menlo Park, CA: Author (2011).
49. J. R. Fraenkel, N. E. Wallen, H. H. Hyun, *How to Design and Evaluate Research in Education* (McGraw-Hill Humanities/Social Sciences/Languages, New York, 2011).
50. C. T. Williams, E. M. Walter, C. Henderson, A. L. Beach, Describing undergraduate STEM teaching practices: A comparison of instructor self-report instruments. *Int. J. STEM Educ.* 2, 18 (2015).
51. Y. Rosseel, Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *J. Stat. Softw.* 48, 1–36 (2012)
52. M. Stains, J. Harshman, M. K. Barker, S. V. Chasteen, R. Cole, S. E. DeChenne-Peters, M. K. Eagan, J. M. Esson, J. K. Knight, F. A. Laski, M. Levis-Fitzgerald, Anatomy of STEM teaching in North American universities. *Science*, 359, 1468–1470, (2018).
53. K. L. Gwet, *How to Compute Intraclass Correlation With MS EXCEL: A Practical Guide to Inter-Rater Reliability Assessment for Quantitative Data* (Advanced Analytics LLC, Gaithersburg, MD, 2010).
54. D. Mogashana, J. M. Case, D. Marshall, What do student learning inventories really measure? A critical analysis of students' responses to the approaches to learning and studying inventory. *Stud. Higher Educ.* 37, 783–792 (2012).

55. A. Stes, L. Coertjens, P. Van Petegem, Instructional development in higher education: Impact on teachers' teaching behaviour as perceived by students. *Instruct. Sci.*, 41, 1103–1126 (2013).
56. J. Nigh, K. E. Pytash, R. E. Ferdig, W. Merchant, Investigating the potential of MOOCs in K-12 teaching and learning environments. *J. Online Learn. Res.* 1, 85–106 (2015).
57. K. Fosnacht, S. Sarraf, E. Howe, L. K. Peck, How important are high response rates for college surveys? *Rev. Higher Educ.*, 40, 245–265 (2017).
58. J. M. Maher, J. C. Markey, D. Ebert-May, The other half of the story: effect size analysis in quantitative research. *CBE Life Sci. Educ.*, 12, 345–351 (2013).
59. R Foundation for Statistical Computing, 161 Vienna, Austria, www.R-project.org/.
60. L. MacNell, A. Driscoll, A. N. Hunt, What's in a name: Exposing gender bias in student ratings of teaching. *Innov. Higher Educ.* 40, 291–303 (2015).