The Role of Syllabi in Measuring Teaching: Comparing Planned and Enacted Behavior

Tim Archie
U. Colorado Boulder

Sarah B. Wise
U. Colorado Boulder

Jenny Valadez Fraire
U. Colorado Boulder

Sandra L. Laursen
U. Colorado Boulder

Timothy J. Weston
U. Colorado Boulder

Syllabi contain information about an instructor's teaching and have been used as a measure of change in response to professional development. Syllabus scores have been linked to more widely used and resource intensive measures of change, including observations and surveys. We used the theory of planned behavior to examine the relationship between syllabus scores and self reported teaching practices and tested for changes in these measures in response to professional development. We found syllabi scores can approximate instructor teaching intentions, which seemed to manifest in their use of teaching practices. Instructors' teaching intentions changed in response to PD, but their teaching practices did not. These findings are consistent with the theory of planned behavior, indicating that intentions change before behaviors, which supports the use of syllabi as a measure of change that can be economically deployed and detect nascent changes in teaching practices.

Keywords: syllabus, teaching, professional development, theory of planned behavior

Non-didactic, student-centered teaching practices are well documented as effective for enhancing student learning and achievement in undergraduate STEM education (Freeman et al., 2014). Evidence also indicates that student-centered teaching practices contribute to reducing achievement gaps among students historically underrepresented in STEM (Ernst & Colthorpe, 2007; Haak et al., 2011; Eddy & Hogan, 2014; Freeman et al., 2014). Despite these benefits, widespread adoption remains limited; estimates suggest that only about 20% of STEM instructors consistently implement student-centered teaching practices in their teaching (Stains et al., 2018; Eagan, 2016). Consequently, many students have little exposure to student-centered teaching practices and miss out on their potential advantages.

Professional development (PD) has been identified as an influential strategy in increasing instructors' use of student-centered teaching practices in undergraduate STEM education (Khatri et al., 2013; Laursen et al., 2019). However, relatively few studies have examined the relationship between teaching-focused PD and the instructional practices of STEM faculty across disciplines such as math, chemistry, physics, and biology (Archie et al., 2022; Benabentos et al., 2020; Borda et al., 2020; Chasteen & Chattergoon, 2020; Manduca et al., 2017; Bathgate et al., 2019). Research in this area may be limited, at least in part, by the research methods commonly utilized in this research, including classroom observations, surveys, and interviews. These research methods are time-consuming and labor-intensive and thus make this type of research expensive to conduct. More economical research methods have the potential to encourage new research, especially given the current funding constraints in STEM education research.

Syllabi are a rich data source that can be collected easily and analyzed quickly using syllabus scoring tools such as the Measuring the Promise (MtP), a syllabus rubric that been shown to detect changes in teaching practices (Mumpower et al., 2024), including changes in response to professional development (CTE, 2017; Favre et al. 2021; Schmid et al., 2021). While the MtP authors emphasize that the "syllabus is only a proxy to actual classroom practices and student

learning and is in itself neither a measure of teaching effectiveness nor necessarily an accurate reflection of an instructor's values," it can be argued that the syllabus can reflect an instructor's intentions for teaching: it is created before an instructor performs any teaching and describes the intent of the course and what student will learn (Palmer et al, 2014a). Research has shown that syllabi and classroom observations are positively associated (Laursen & Archie, 2018), suggesting that instructor intent indeed translates to their teaching.

The temporal relationship between syllabi and teaching practices is well aligned with the Theory of Planned Behavior (TPB) in which behavioral intentions are the primary determinants of actual behavior (Ajzen, 1991). Intentions, in turn, are shaped by three key constructs: attitude, subjective norms, and perceived behavioral control. Attitude refers to an individual's overall positive or negative evaluation of a given behavior. Subjective norms represent an individual's perception of social pressure to carry out or avoid the behavior. Perceived behavioral control reflects the extent to which an individual believes they can perform the behavior and exercise control over it. Importantly, for intentions to translate into action, individuals must possess the requisite ability to perform the behavior. Consequently, perceived behavioral control is hypothesized to exert a direct influence on actual behavior, in addition to its effect on intentions.

The TPB has been employed to examine shifts in teaching and learning within STEM education (Reinholz & Andrews, 2020). Specifically, it has been used to explain mathematics and physics instructors' adoption of research-based instructional strategies after participating in professional development programs (Archie et al., 2022; Chasteen & Chattergoon, 2020). In this study, the TPB is used to explore the relationship between the syllabus and instructors' teaching practice, using the MtP syllabus rubric as an indicator of change in practice. In this study, we answered two research questions:

1. What is the relationship between instructors' behavioral intent (syllabi scores) and their behavior (use of student-centered teaching scores)?
2. Do instructors' behavioral intentions and behavior change after participating in professional development?

Methods

The study design was approved by the university's IRB and all instructors gave informed consent.

Study Population and Survey Sample

Study participants were early-career undergraduate mathematics instructors who participated in MAA Project NExT), New Experiences in Teaching, professional development program. Briefly, this full-year program provides strategies and advice for teaching and navigating the academic environment, with an emphasis on a variety of inclusive and student-centered approaches to teaching (MAA, n.d.; Higgins, 2012). Because this analysis focuses on syllabi as a measure of change, rather than on outcomes of the PD, we do not describe the program in detail.

Participants completed a pre-survey at the end of the academic term preceding participation in the PD program, a mid point survey 1 year into the PD program, and a final survey at the end of the PD program. The survey asks respondents to report on a specific course they taught and to upload the syllabus for their selected course. A total of 171 instructors completed the pre survey and provided a syllabus, 110 completed the midpoint survey and provided a syllabus, and 72 completed a final survey and provided a syllabus. To maximize statistical power, we computed a "post" syllabi score from the final time point (n = 72) or the mid time point for survey

participants who provided a syllabus at the mid time point, but not at the final time point (n = 70), for a total of 142 respondents.

Of the 171 unique respondents, 168 provided some demographic information. 48% identified as men, 46% identified as women, 5% identified as other than man or woman. 72% identified as white, 15% identified as Asian, 8% identified as Hispanic, and 3% identified as Black. 61% were assistant professors and 34% were non-tenure-track instructors or visiting professors. 41% taught at a PhD granting institution, 44% at a 4-year institution, and 15% at a Master's granting institution. 16% taught at a minority-serving institution, and the remaining 84% did not or did not know if they taught at a minority-serving institution.

Survey measures. In this study, the TPB construct of "behavior" was operationalized using self-reported teaching practices. Respondents were asked to report the percentage of time they used 11 teaching practices in a specific course over the course of the term, totaling to 100%. We then computed "behavior": as the proportion of time using three student-centered teaching practices (student group work + student presentation + class discussion). Syllabi scores were used as a measure of the TPB construct "behavioral intent", and the scoring procedure is described in the following section. Departmental expectations to use non-lecture based teaching strategies represented the "subjective norm" construct. The item was worded as "How much do colleagues in your department have formal or informal expectations for instructors' teaching? - Expectation to use a variety of techniques" and was coded on a five-point scale (1=not at all, 5=a great deal).

The TPB construct, "perceived behavioral control", was operationalized by one item, "How much effort have you made to implement active learning strategies in your teaching?" and was coded on a five-point scale (1=no effort, 5= a great deal of effort).

Instructors were asked three growth mindset questions (Dweck, 2006) that have been shown to be related to the use of student-centered teaching practices (Aragón et al., 2018). These items were reverse coded on a five-point scale (1 = strongly agree, 5 = strongly disagree).

Syllabus scoring. Prior to coding, syllabi were deidentified with respect to instructor name(s) and contact info, higher education institution, and other wording that might identify the course. Calendar year dates were removed to mask whether the syllabus was part of the pre-program or follow-up sample.

Two of the authors (SL, SW) followed an iterative process to test and modify the Making the Promise (MtP) rubric. They first applied the MtP rubric to four annotated reference syllabi from the rubric authors (CTE, 2017) and discussed findings. They then applied the MtP rubric to 15 college-level course syllabi collected in prior mathematics education studies. The authors scored several syllabi individually, then discussed their scoring rationales until they came to consensus. They revised the rubric to specify the levels of evidence for each item, to clarify rubric language, and to distinguish items so they were non-overlapping, and repeated this process several times across the set of 15 syllabi. This yielded a detailed modified rubric outlining three scoring levels (low, medium, and high) for each of 15 items. An overview of the modified MtP rubric items and scoring is provided in Figure 1.

Like the original MtP rubric, our modified rubric contained five subsections: learning goals, assessment, schedule, classroom environment, and class activities. Added language specifies the kinds of evidence that support scoring and allows the rater to reliably assess whether the strength of evidence for each item is low, medium, or high. These descriptions also support the rater in making decisions around evidence that can be inferred from the syllabus text. Such evidence can be used to support scoring an item as medium, but cannot be used to support a high score.

Following this iterative process, the two authors individually rated a new set of ten mathematics syllabi, five collected for this study. Author TW calculated interrater reliability statistics on the results, finding an interrater agreement of 65%, an intraclass correlation (ICC) of 0.73, and a Kappa of 0.67. These statistics are in acceptable ranges for interrater reliability; a Kappa of 0.618-0.8 is considered "substantial" (Cohen et al., 2000). One rater subsequently applied the rubric to the entire dataset of mathematics syllabi (n = 380). Eight syllabi were excluded due to substantial missing information (e.g., grading scheme) that made the syllabus impossible to score, or a lack of alignment with inclusion criteria (e.g., not an undergraduate course). A total of 353 syllabi were scored on each of 15 individual items. Low, medium, and high item scores were converted to 0, 1, or 2 points, respectively, then multiplied by the item weight (Figure 1) to arrive at a numerical score. We adjusted the weighting of items #9 and #14, as described in Figure 1. The total maximum score for each syllabus was 54.

Any rubric is potentially influenced by bias, and bias can be magnified by differential weighting of items. We reduced item weighting where our syllabus set tended to lack information important to evaluating the item, but did not alter the weighting of the original authors' judgments of the relative importance of different items. Overall, the original MtP rubric puts significant weight on learning goals, learning objectives, and the alignment of those goals with assessments and classroom activities, and less heavily weights the quality of classroom activities and the classroom environment. A differently weighted rubric could generate score patterns and analyses with slightly different implications.

Item	Short definition	Modifications made	Item weight	Max item score
1	Big picture learning goals and measurable smaller learning objectives are articulated or implied. Together, goals and objectives address all of Fink's six dimensions of learning (Fink, 2013).	Focused on breadth of learning goals, operationalized as number of Fink's dimensions included in the set of learning goals and objectives (rather than mere presence of learning goals).	3	6
2	Learning goals and learning objectives are separately defined and clearly articulated. Learning goals be may written broadly, but learning objectives are measurable and use specific action verbs.	Reduced overlap with item #1 with a focus on the degree to which learning goals are clearly and measurably articulated (rather than the presence of learning objectives).	3	6
3	Learning objectives are appropriately pitched to the course level, class size, position of the course within the curriculum, and characteristics of students taking the class.	Not assessed for this study, due to a lack of needed contextual information.	0	0
4	Assessments are aligned with learning goals and objectives.	Focused on alignment only; removed considerations of grade weighting.	3	6
5	Major assessments are well defined, rationalized, and involve "authentic tasks".	No changes.	2	4
6	There is frequent formative assessment and scaffolding of learning that encourages students to attend to past learning and iterate their efforts.	Added wording about scaffolding and iteration of effort.	1	2
7	Assessments are adequately paced, spreading out the student workload and giving students information in a timely manner to make decisions about enrollment.	Focuses on how well pacing information aids students in making enrollment decisions. Removed wording about scaffolding, reducing overlap with #6.	1	2
8	The grading scheme is weighted appropriately for the assessment design.	Focus is placed on grade weighting. Removed wording about goals alignment. Both changes reduce overlap with #4. Removed wording about where the grading scheme is placed in the syllabus.	1	2
9	Schedule is fully articulated and provides evidence for a logical progression of learning.	Reduced weighting from 3 to 2 because schedule information was often provided in a course LMS, not accessible to this study	2	4
10	Invites the learner and indicates a positive and collaborative class environment.	Removed wording about "big questions" addressed by course content.	2	4

11	Supports connected and lifelong learning. Answers the question: Why should you learn this for your life?	Removed considerations about grading schemes to reduce overlap with #8.	2	4
12	Signposts that the work will be challenging, and students will be supported to enable their success.	Focused on methods for supporting students through challenging work. Removed wording about rigor and expectations to reduce overlap with #1, 2, 4, and 8.	1	2
13	An organized, learning-supportive, interactive document.	No changes.	1	2
14	Classroom activities are aligned with assessments.	Removed wording about alignment with learning objectives, to reduce overlap with #4. Reduced score weighting from 3 to 2 because information needed to score this item was rarely provided.	2	4
15	Learning activities are derived from evidence-based practices.	No changes.	2	4
16	Class activities engage students in a variety of ways.	No changes.	1	2

Figure 1. Rubric items, definitions, and modifications to the original Making the Promise rubric and scoring.

Data Analysis

A structural equation model was used to answer RQ1. (AMOS 31, SPSS Inc., Chicago, IL, USA). This analysis was limited to 99 survey responses with complete data and only included one response from any individual respondent across time points. The model was calculated using maximum likelihood estimation and model fit was evaluated using two indices: the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA). CFI values range from 0 to 1, with values ≥ 0.95 indicating good fit (Hu & Bentler, 1999). RMSEA values ≤ 0.05 are interpreted as evidence of close fit (Browne, 1993).

RQ2 was answered using independent samples t-tests to test for differences in pre and post syllabus scores and average use of student-centered teaching practices.

Results

A path model was used to answer RQ1. The model fit indices were within acceptable ranges (CMIN = (df) 3.830 (2) $p = 0.147$, CFI = 0.99, RMSEA =0.035). As shown in Figure 2, we found positive, statistically significant standardized direct effects in all but one of the specified paths in the structural model. Among the four variables hypothesized to influence behavioral intention (syllabi scores), both perceived behavioral control (effort in implementing student-centered practices) and attitude (growth mindset items) showed comparable moderate effects, while subjective norm (departmental expectations for teaching) showed a minimal and non-significant direct effect. Behavioral intention and perceived behavioral control were both expected to directly influence behavior (use of student-centered teaching practices), with each showing a moderate positive association with behavior. The model explained moderate amounts of variability in both behavioral intention ($R^2 = 0.23$) and behavior ($R^2 = 0.31$).

T-tests used to answer RQ2 revealed statistically significant positive change (Table 1) in overall syllabi scores and in two of the five syllabus categories: learning goals and assessments. All other categories showed positive minimal effects that were not statistically significant. Instructor use of student-centered teaching practices shifted from 32% to 35% from pre to post, a minimal and not statistically significant difference.

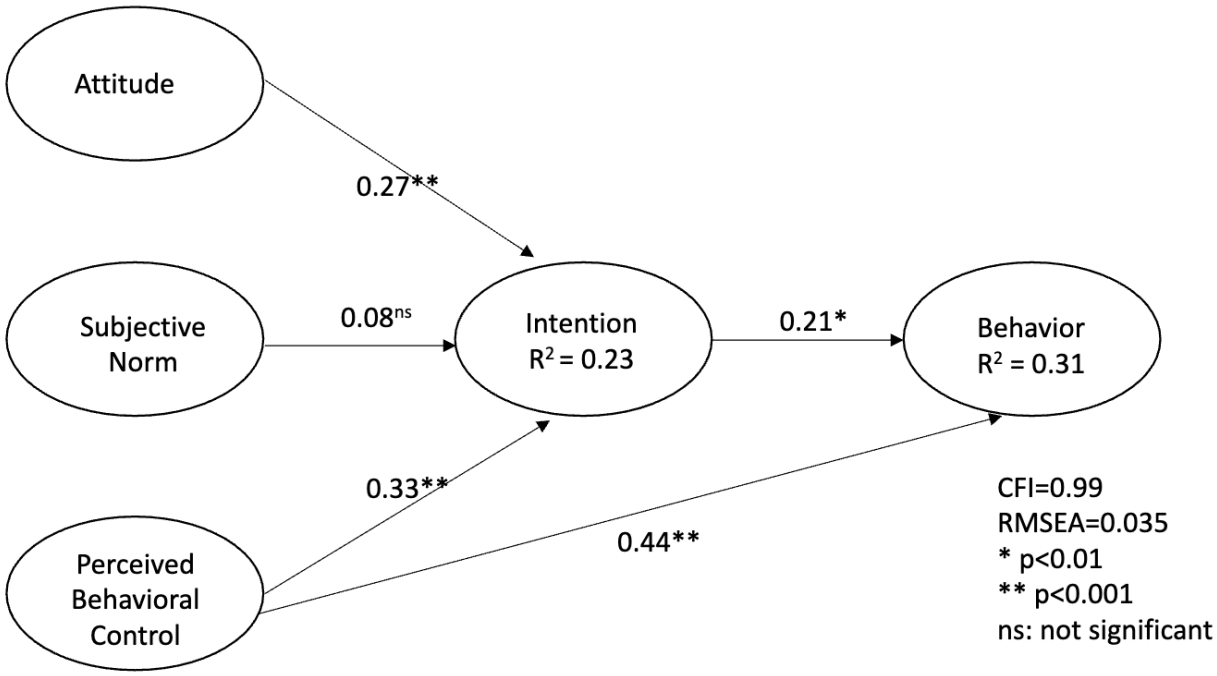

Figure 2. Structural equation model of the theory of planned behavior with standardized direct effects.

Table 1. Pre/post comparison of syllabus scores, subscores, and average percent course time using student-centered teaching practices: T-test results

Category	Pre	Post	*t*	*df*	*p*	*d*
Overall score	24.3	26.7	2.724	324	0.007	0.31
Learning goals	5.98	7.12	3.300	324	0.001	0.37
Assessments	8.73	9.58	3.100	324	0.002	0.35
Schedule	1.70	1.65	0.455	324	0.650	0.05
Environment	4.37	4.51	0.511	324	0.610	0.06
Learning activities	3.52	3.83	0.961	324	0.337	0.11
Teaching practices	32.35	34.55	0.792	324	0.429	0.09

Discussion

RQ1 results showed that syllabi scores are positively associated with instructors' use of student-centered teaching practices and are consistent with the TPB and prior research. These findings indicate that instructors' teaching is aligned with their teaching intentions as communicated in their syllabi and is consistent with prior TPB research, which has found behavioral intent to be a strong predictor of behavior (Armitage & Connor, 2001). Also consistent with prior research, we observed that behavioral intent and behavior were moderately influenced by perceived behavioral control. We found that the greater the effort of instructors to use active learning teaching practices, the more likely their syllabus and teaching were student-centered. Attitudes about students also showed a positive relationship with syllabus scores, indicating that an instructor's attitude influenced their intentions for the course and, subsequently, their teaching practices. Lastly, we found no significant association between subjective norm and behavioral intention, indicating that perceived departmental expectations for using a variety of teaching practices did not relate to instructors' syllabus scores. While this finding is unexpected and inconsistent with some prior research (McConnell et al., 2020,

Bathgate et al., 2019), it is consistent with prior applications of the TPB that have also found no meaningful association between subjective norms and behavioral intent (Armitage & Conner, 2001; Archie et al, 2022).

For RQ2, the results showed statistically significant positive change in overall syllabi scores and in two of the five syllabus categories: learning goals and assessments. All other categories showed positive but minimal effects that were not statistically significant. These results suggest that the overall shift in the syllabi scores was driven primarily by these two categories. It is plausible that some PD content was directly or indirectly related to articulating their learning goals in their course materials. The lack of meaningful change from pre to post average use of student-centered teaching practices was unexpected. However, because much of the data used in this study was collected during the pandemic, these findings are not likely to be fully reflective of their teaching constrained by the challenge of adapting student-centered teaching practices to remote teaching formats.

The findings from this study are encouraging as it provides evidence that syllabi can be a useful measure of an instructor's teaching practice. While the syllabus may not reflect enacted teaching practices, our findings indicate that a syllabus scored with the MtP rubric can approximate the type of teaching practices used by a given instructor. A syllabus can be a more accessible and cost-effective means of characterizing teaching and measuring differences across time points in response to teaching-focused interventions such as professional development.

While the findings from RQ1 indicate the intention that instructors infuse into their syllabi manifests in their teaching behaviors (and use of student-centered teaching practices), the findings from RQ2 showed a positive shift in intention, but no significant change in teaching practices. These findings suggest that while instructors intend to change their teaching practices, they may need more time to gain skills and knowledge to use the teaching practices they aspire to use. Thus, the syllabi and corresponding MtP score may be a more useful and realistic measure of short-term PD outcomes than more resource-intensive methods such as observations or surveys.

Acknowledgments

This work is supported by the National Science Foundation award DUE-2111273. The opinions, findings, and conclusions expressed here are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Anatomy of STEM teaching in North American universities. Science, 359, 1468-1470.

Ajzen, I. (1991). The theory of planned behavior. *Organizational and Human Decision Processes*, 50(2), 179-211.

Archie, T., Hayward, C. N., Yoshinobu, S., & Laursen, S. L. (2022). Investigating the linkage between professional development and mathematics instructors' use of teaching practices using the theory of planned behavior. *Plos one*, 17(4), e0267097.

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British journal of social psychology*, 40(4), 471-499.

Bathgate, M. E., Aragón, O. R., Cavanagh, A. J., Waterhouse, J. K., Frederick, J., & Graham, M. J. (2019). Perceived supports and evidence-based teaching in college STEM. *International Journal of STEM Education*, 6(1), 1-14.

Benabentos, R., Hazari, Z., Stanford, J. S., Potvin, G., Marsteller, P., Thompson, K. V., Cassone, V. M., Murasko, D., & Kramer, L. (2020). Measuring the implementation of student-centered

teaching strategies in lower-and upper-division STEM courses. *Journal of Geoscience Education*, https://doi.org/10.1080/10899995.2020.1768005

Borda, E., Schumacher, E., Hanley, D., Geary, E., Warren, S., Ipsen, C., & Stredicke, L. (2020). Initial implementation of active learning strategies in large, lecture STEM courses: Lessons learned from a multi-institutional, interdisciplinary STEM faculty development program. *International Journal of STEM Education*, 7(1), 1-18.

Browne, M. (1993). Alternative Ways of Assessing Model Fit. Testing Structural Equation Models, 136-162.

Center for Teaching Excellence (2017). Syllabus Rubric. University of Virginia. http://cte.virginia.edu/resources/syllabus-rubric/

Chasteen, S. V., & Chattergoon, R. (2020). Insights from the Physics and Astronomy New Faculty Workshop: How do new physics faculty teach?. *Physical Review Physics Education Research*, 16(2), 020164.

Cohen, J. A., Fischer, J. S., Bolibrush, D. M., Jak, A. J., Kniker, J. E., Mertz, L. A., ... & Cutter, G. R. (2000). Intrarater and interrater reliability of the MS functional composite outcome measure. *Neurology*, 54(4), 802-806.

Eagan, K. (2016). Becoming more student-centered? An examination of faculty teaching practices across STEM and non-STEM disciplines between 2004 and 2014. Alfred P. Sloan Foundation.

Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work?. *CBE—Life Sciences Education*, 13(3), 453-468.

Ernst, H., & Colthorpe, K. (2007). The efficacy of interactive lecturing for students with diverse science backgrounds. *Advances in physiology education*, 31(1), 41-44.

Favre, D. E., Bach, D., & Wheeler, L. B. (2021). Measuring institutional transformation: a multifaceted assessment of a new faculty development program. *Journal of Research in Innovative Teaching & Learning,* 14(3), 378-398.

Fink, L. D. (2013). *Creating significant learning experiences: An integrated approach to designing college courses*. John Wiley & Sons.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410-8415.

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science*, 332(6034), 1213-1216.

Higgins, A. (2012). Project NExT: A program of the Mathematical Association of America. *In Council of Scientific Society Presidents (CSSP), The role of scientific societies in STEM faculty workshops: A report of the May 3, 2012 meeting*, pp. 56-62. American Chemical Society. https://www.aapt.org/Conferences/newfaculty/upload/CSSP_May_3_Report_Final_Final_Version_3-15-13.pdf

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: A multidisciplinary journal*, 6(1), 1-55.

Khatri, R., Henderson, C., Cole, R., & Froyd, J. (2013, January). Successful propagation of educational innovations: Viewpoints from principal investigators and program. *In AIP Conference Proceedings* (Vol. 1513, No. 1, pp. 218-221). American Institute of Physics.

Laursen, S., & Archie, T. (2018). How do we teach thee? Let me count the ways. A syllabus rubric with practical promise for characterizing mathematics teaching. In A. Weinberg, C. Rasmussen, J. Rabin, M. Wawro, & S. Brown (Eds.), *Proceedings of the 21st Annual Conference on the Research in Undergraduate Mathematics Education* (pp. 862-870). San Diego, CA: Mathematical Association of America, SIGMAA on RUME.

Laursen, S., Andrews, T., Stains, M., Finelli, C. J., Borrego, M., McConnell, D., Johnson, E., Foote, K., Ruedi, B., & Malcom, S. (2019*). Levers for change: An assessment of progress on changing STEM instruction*. Washington, DC: American Association for the Advancement of Science.

MAA (no date). MAA Project NExT. https://maa.org/maa-project-next/

Manduca, C. A., Iverson, E. R., Luxenberg, M., Macdonald, R. H., McConnell, D. A., Mogk, D. W., & Tewksbury, B. J. (2017). Improving undergraduate STEM education: The efficacy of discipline-based professional development. *Science Advances*, 3(2), e1600193.

McConnell, M., Montplaisir, L. & Offerdahl, E. (2020). A model of peer effects on instructor innovation adoption. *International Journal of STEM Education*, 7, 53. https://doi.org/10.1186/s40594-020-00255-y

Mumpower, L. A., Rohrbacher, C., Caulkins, J., & Korentsides, J. (2024). Measuring pandemic change: Analyzing syllabi for student-centeredness over time. *To Improve the Academy: A Journal of Educational Development*, 43(2).

Palmer, M., Bach, D., & Streifer, A. (2014a). Measuring the Promise: A Valid and Reliable Syllabus Rubric. Guide to Assessing the Focus of Syllabi. University of Virginia Teaching Resource Center.

Palmer, M. S., Bach, D. J., & Streifer, A. C. (2014b). Measuring the promise: A learning-focused syllabus rubric. *To Improve the Academy*, 33, 14-36. doi:10.1002/tia2.20004

Schmid, M. E., Bajcz, A. W., & Balster, N. J. (2021). Evaluating a novel faculty development program in teaching at a research-intensive university. *Teacher Development*, 25(3), 340-365.

Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., ... & Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science*, *359*(6383), 1468-1470.