

Investigation of physics departments' assessment and change practices

Nathalia S. Martinez Garcia, Chandra Turpen, and Robert P. Dalka
Department of Physics, University of Maryland, College Park, MD 20742

Joel C. Corbo
Center for STEM Learning, University of Colorado Boulder, Boulder, CO 80309

It is important that change in physics programs is deliberate, evidenced-based, and engages multiple stakeholders. To assess the state of departmental change practices, the Effective Practices for Physics Programs (EP3) Initiative regularly runs a survey of department chairs to measure departmental cultures of assessment. The 2023 survey received 163 research-consenting responses. This paper presents two claims based on these survey results: (1) assessments are largely not seen as leading to change, although chairs aspire for them to do so, and (2) chairs see substantial room for improvement in how they go about changing the undergraduate physics program, especially when it comes to engaging multiple stakeholders and using data effectively. The significant difference between current and ideal points to areas where shifting the culture within departments could have support from departmental leadership.

I. INTRODUCTION

It is important that change in physics programs is guided by lessons that have been learned over the years of research into higher education organizational change. In a systematic review of change literature in higher education, Kezar reported that unsuccessful change efforts are driven by tacit theories of change, ignore the specific contexts, assume simplistic change models, and are not grounded in the research [1]. To support physics faculty members in avoiding these mistakes, the Effective Practices for Physics Programs (EP3) Initiative was launched by the American Physical Society (APS) and the American Association of Physics Teachers (AAPT). The EP3 Initiative aims to help physics programs respond to challenges they face by gathering and sharing knowledge, experience, and proven good practice. This is done both through professional development activities to support physics faculty change leaders and resources such as the EP3 guide. The EP3 Guide was developed for self-assessment of undergraduate physics programs and how to pursue change within them [2].

The philosophy of the EP3 Initiative is that change is deliberate, based in evidence, context-dependent, collaborative, and ongoing. Similar to continuous improvement frameworks in higher education, this approach to change focuses on a commitment to stakeholder needs/concerns/values, ongoing self-improvement, and ongoing self-assessment [3–5]. Establishing a culture of assessment is an important part of this change process. A culture of assessment encapsulates the cultural norms and contexts pertaining to use of assessment data to support undergraduate education [6]. In contrast to a culture of compliance—where assessment is done just to fulfill external requirements—a culture of assessment embraces things such as program reviews to improve departments and programs [7].

In 2020, the EP3 Evaluator and Research Team surveyed department chairs to investigate many different departmental factors that were related to change work. This included (a) understanding threats to program sustainability, (b) specific problems departments faced, (c) the departmental culture of assessment and change, and (d) the approaches to program review [2]. In this paper, we present findings from a follow up survey administered in 2023, specifically focused on documenting departmental culture of assessment and change. We are guided by the following research questions: (1) how are assessments used in physics departments when pursuing change? and (2) what practices are a part of departmental change strategies?

II. METHODS

This quantitative study used a survey instrument designed to measure departmental cultures of assessment. We sought to measure cultures of assessment within physics departments across the country at different points in time, particularly fol-

lowing the release of the EP3 Guide [2] and its accompanying community engagement activities. Because it was not feasible with our resources to get the perspectives of a broad array of departmental stakeholders, we decided to survey department chairs as a proxy for the state of departments as a whole. While this limits our ability to understand the full scope of departmental culture, it gives us an understanding of how departmental leadership views their assessment culture. The survey instrument used in this study was adapted from prior survey instruments [6–8] and was administered in collaboration with the APS, which maintains an email list of physics department chairs.

The survey was sent to 734 unique institutions with physics programs on June 7, 2023 and remained open until Nov. 3, 2023. We received 188 complete responses to the items considered here; however, 25 respondents did not consent to having their responses used for research purposes. Table I provides a breakdown of the 734 unique institutions that received the survey and the 162 that completed the survey and consented to the use of their data in this research study. This results in a 22.1% response rate (for respondents included in this analysis). This breakdown includes highest physics degree type awarded (Bachelor’s, Master’s, Doctoral), public vs. private, and Minority Serving Institution (MSI) vs. non-MSI. Across our sample, we found that: (a) non-MSIs are over-represented while MSIs are under-represented and (b) the balance of public and private institutions and across degree-granting institutions is fairly representative. Specific types of MSIs that are underrepresented in this data set include Asian American, Native American, and Pacific Islander Serving Institutions (AANAPISIs), Alaska Native and Native Hawaiian Serving Institutions (ANNHIs), Historically Black Colleges and Universities (HBCUs), and Hispanic Serving Institutions (HSIs).

The survey framed "assessment of student learning" as including assessment of program-level learning outcomes, but not including grades or drop-fail-withdraw rates. The survey framed "Internally-analyzed departmental outcomes assess-

Inst. Characteristic	% of inst. with characteristic out of all inst. that received the survey (N=734)	% of inst. with characteristic out of all inst. in the sample of respondents (N=162)
Bachelor’s granting	66% (N=483)	68.5% (N=111)
Master’s granting	7% (N=51)	8.0% (N=13)
Doctoral granting	27% (N=200)	23.5% (N=38)
Public	49.2% (N=361)	48.1% (N=78)
Private	50.8% (N=373)	51.9% (N=84)
MSI	21.3% (N=156)	11.7% (N=19)
Non-MSI	78.7% (N=578)	88.3% (N=143)

TABLE I. We compare our sample of respondents from unique institutions (in the right hand column) to the landscape of all unique institutions that award degrees in physics (in the left hand column).

ment" as including the assessment of departmental outcomes other than student learning, including (but not limited to): Recruitment, Retention, Grades, Drop-fail-withdraw rates, Department climate, Satisfaction surveys, Exit interviews, Focus groups, Alumni feedback.

All survey items considered here were posed on a seven-point response scale of 'strongly disagree' to 'strongly agree.' Each survey item asked respondents to consider the item statement in relation to their typical "current" departmental practices AND in relation to their perceptions of "ideal" departmental practices. So for each item we have a pair of responses to compare for each respondent. These surveys can help us deepen our understanding of a department's culture. It is important to ask individuals about their current perceptions as well as what they would consider to be ideal qualities of an undergraduate program. Focusing on current and ideal aspects of departments allow our survey to function as a tool for formative feedback to department members, because it illustrates what they aspire to be, as well as where the department perceives it to be falling short [8].

To test the significance of differences between respondents' "ideal" and "current" paired responses to an item, we use a Wilcoxon signed rank test, which is appropriate for paired-responses of ordinal non-parametric data [9]. The outputs of this test include the W s test statistic, a z -score, and a p -value. All items were statistically significant at the $p < 0.0001$ level. We interpret the difference between the respondents' "current" response and their "ideal" response as a gap between the current state of their department and their perceived ideal state for their department [8].

III. RESULTS

Our results from the EP3 survey of department chairs substantiated two central claims: (1) assessments are largely not seen as leading to change, although chairs aspire for them to do so, and (2) chairs see substantial room for improvement in how they go about changing the undergraduate physics program, especially when it comes to engaging multiple stakeholders and using data effectively. Our analysis relies on 6 items to support claim (1) and 7 items to support claim (2).

A. Assessments are largely not seen as leading to change, although chairs aspire for them to do so.

The survey items in Table II for claim (1) are based on how a department views of assessments practices and their relationship to change efforts.

To investigate how departmental assessments inform programmatic change, we analyzed survey questions that asked respondents to think about both perceptions and actions related to assessment and change. Between 81% and 88% of respondents agreed or strongly agreed (excluding slightly agree) that their ideal departments would have departmental

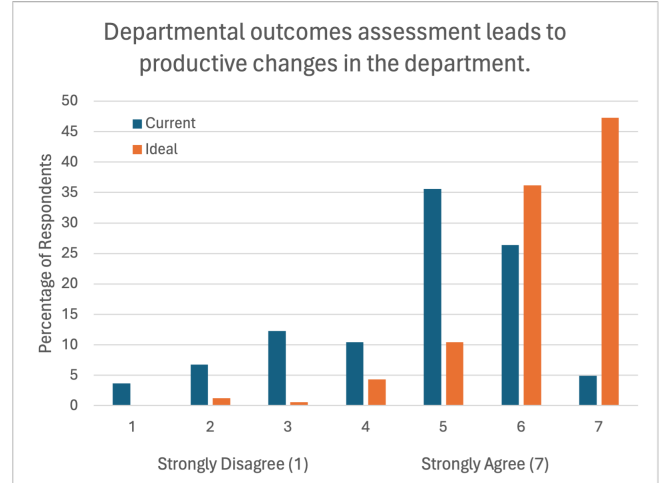


FIG. 1. Histogram featuring both current (blue) and ideal (orange) responses to the survey question "Internally-analyzed departmental outcomes assessment leads to productive changes in the department." It is chosen as it is representative of the distributions for all of the survey items of this claim.

outcome assessments that occur at least once a year, were perceived by faculty as a valuable driver of change, and lead to productive changes in their departments. However, only 63% of respondents agreed that such assessments occurred at least once a year (a gap of 25%). Only 31% of respondents agreed that outcomes assessments were perceived by faculty as a valuable driver for change in their current departments making this a 50% gap. Finally, 31% of respondents also agreed that these outcomes assessments lead to productive changes in their current department making this a 52% gap between their current and ideal departments.

Similarly, between 79% and 88% of respondents agreed or strongly agreed that in their ideal departments, assessments of student learning are perceived by faculty as a valuable driver of change, would be perceived by faculty as helping improve student learning, and lead to productive changes in the department. Again, the gaps between current and ideal departments were between 57% and 62% meaning that departments are far from their ideal for their departments.

These gaps are exemplified in Figure 1, which highlights the survey item focused on assessments leading to productive changes and was chosen as it was the median gap represented in this claim. Figure 1 shows that the current histogram peaks at 5 whereas the ideal histogram peaks at 7, showing more respondents agreeing with their ideal rather than their current department.

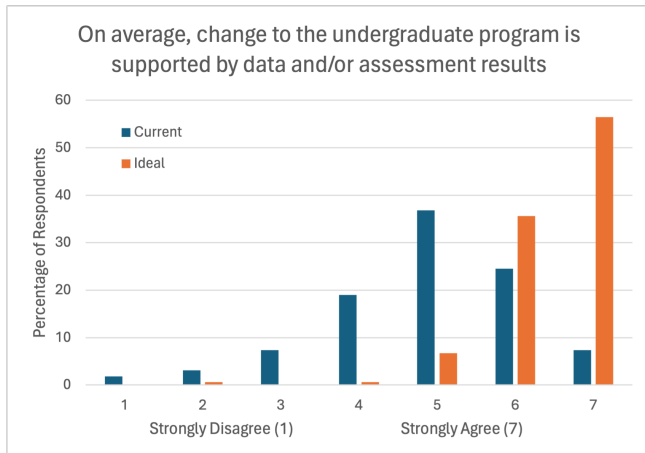


FIG. 2. Histogram featuring both current (blue) and ideal (orange) responses to the survey question “On average, change to the undergraduate program is supported by the data and/or assessment results.” It is chosen to illustrate an item with a more substantial gap between current and ideal.

B. Chairs see substantial room for improvement in how they go about changing the undergraduate physics program, especially when it comes to engaging multiple stakeholders and using data effectively.

The survey items in Table II for claim (2) are based on how changes to a department are being implemented.

To investigate undergraduate programmatic change in departments, we analyzed survey questions that asked respondents to report their current and ideal department’s engagement with a range of change practices. Only 25% of the respondents agreed or strongly agreed (excluding slightly agree) that changes to the undergraduate program attempted to involve multiple stakeholders. Only 32% agreed that they are supported by data and/or assessment results. However, between 50% and 62% of respondents agreed that their current department is driven by departmental goals or concerns, is driven by a sense of the department’s purpose, is seen as an ongoing process rather than an event, is supported by the existence of a departmental culture, and is driven by a shared responsibility among department faculty for the health of the department and the people in it. Overall, the respondents’ current and ideal departments have a lower gap.

Figure 2 highlights a survey item focused on change supported by data and/or assessment results which was chosen to illustrate a particular high gap between current and ideal departmental practices. The trend seen in Figure 2 is similar to that of Figure 1 as it also shows that the current histogram peaks at 5 whereas the ideal histogram shows progressively larger values from 5 to 7. Most of the respondents believed that their current department’s engagement with change practices falls substantially short of the ideal. As seen in Table II, the gap between respondents’ current and ideal departments was larger for claim (1) than it was for claim (2).

IV. CONCLUSION

The current practices around data use indicate that many departments operate under a culture of compliance, where assessment of their programs and students is not used to inform meaningful changes in their programs, but is rather done to meet external requirements. Even if departments engaged in a change effort, without substantial shifts in their approach to assessment, they could fall into the common mistakes outlined by Kezar [1]. For change efforts within departments to be sustained, it is important that physics departments adopt a culture of assessment. We see that department chairs report that these practices would be a part of their ideal departments. This significant difference between current and ideal points to areas where shifting the culture within departments could have support from departmental leadership.

These results are based on responses from a single perspective within each department. In this way, they are limited in telling us about physics departmental culture as a whole. However, the results reported here align with findings that the EP3 research team has reported on in other contexts. For example, in case studies of departments who were involved with EP3-supported professional development, it was found that many departments (prior to the professional development experience) often had a simplistic approach to change that relied on individual efforts without deliberate consideration of data or evidence [10]. Although we only have a single data point for each department in this study, it is clear that leaders in physics programs desire improved change practices, and to move away from the ways change has been approached historically. In other research of EP3 community engagement activities, it has been shown that these new practices are able to be taken up within departments by offering significant support to developing change leaders [11–13].

This nationwide survey of physics departments about departmental assessment and programmatic change provided valuable insights into the practices, structures, and culture of departments in the U.S. These results will be used to guide the EP3 project, and other national efforts in physics and will likely be administered again in the future to document how practices in physics departments change over time.

ACKNOWLEDGMENTS

We thank all of the survey participants for their time. We also thank the members of the EP3 research team not listed as authors here—Diana Sachmpazidi and Fatima Abdurrahman—for their insights and discussion. Additionally, we thank EP3 Project Membership Team members including, Micheal Wittmann, Michael Jackson, Stephanie Chasteen, Sam McKagan, Kathyne Woodle, Christine O’Donell, Sean Costello, Theodore Hodapp, Robert Hilborn. This work is supported partially by the NSF #1821372. RPD was supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE 1840340.

Claim 1: <i>Assessments are largely not seen as leading to change, although chairs aspire for them to do so.</i>				
	Item Statement	Current	Ideal	Wilcoxon Signed-Rank Test
Internally-analyzed departmental outcomes assessment...	occurs at least once a year, for some element of the department.	63%	88%	$W_s = 159.5$ $p < 0.001$ $z = -6.77$
	is perceived by faculty as a valuable driver of change in the department.	31%	81%	$W_s = 78.0$ $p < 0.001$ $z = -9.46$
	leads to productive changes in the department (e.g. recruitment or retention practices, curricular modifications, etc.).	31%	83%	$W_s = 145.0$ $p < 0.001$ $z = -9.61$
Assessment of student learning...	is perceived by faculty as a valuable driver of change in the department.	22%	79%	$W_s = 127.0$ $p < 0.001$ $z = -9.67$
	is perceived by faculty as helping improve student learning in the department.	30%	87%	$W_s = 119.0$ $p < 0.001$ $z = -9.53$
	leads to productive changes in the department (e.g. how the department prepares students for future careers, curriculum, instruction).	26%	88%	$W_s = 149.5$ $p < 0.001$ $z = -9.92$
	Is [not] primarily done to fulfill institutional or external requirements.	11%	39%	$W_s = 419$ $p < 0.001$ $z = -9.22$
Claim 2: <i>Chairs see substantial room for improvement in how they go about changing the undergraduate physics program, especially when it comes to engaging multiple stakeholders and using data effectively.</i>				
	Item Statement	Current	Ideal	Wilcoxon Signed-Rank Test
On average, change to the undergraduate program...	includes an attempt to involve multiple stakeholders in the department (e.g. students, faculty, advisors, alumni).	25%	83%	$W_s = 26.5$ $p < 0.001$ $z = -9.94$
	is supported by data and/or assessment results.	32%	92%	$W_s = 114.5$ $p < 0.001$ $z = -10.03$
	is driven by departmental goals or concerns (rather than external mandates or threats).	50%	88%	$W_s = 383.0$ $p < 0.001$ $z = -7.69$
	is driven by a sense of the department's purpose (e.g., mission and/or vision).	50%	88%	$W_s = 104.0$ $p < 0.001$ $z = -8.30$
	is seen as an ongoing process rather than an event (e.g., improvement efforts require continued attention).	62%	93%	$W_s = 104.5$ $p < 0.001$ $z = -8.29$
	is supported by the existence of a departmental culture where innovation, experimentation, and learning from past initiatives is encouraged.	58%	94%	$W_s = 247.5$ $p < 0.001$ $z = -8.09$
	is driven by a shared responsibility among department faculty for the health of the department and the people in it.	61%	94%	$W_s = 267.0$ $p < 0.001$ $z = -8.13$

TABLE II. Provides item statements that support claim (1) and claim (2), the percentage of all respondents that answered a 6 or 7 for both the current and ideal departments, and the Wilcoxon signed rank values W_s , p-value, and z-score.

-
- [1] A. Kezar, *How colleges change: Understanding, leading, and enacting change* (Routledge, 2018).
- [2] *Effective practices for physics programs guide*, <https://ep3guide.org/>, accessed: 2024-05-12.
- [3] S. O'Reilly, J. Healy, and R. O'Dubhghaill, *International Journal of Productivity and Performance Management* **67**, 260 (2018).
- [4] C. Temponi, *Quality assurance in education* **13**, 17 (2005).
- [5] K. J. Fryer, J. Antony, and A. Douglas, *The TQM magazine* **19**, 497 (2007).
- [6] M. B. Fuller and S. T. Skidmore, *International journal of educational research* **65**, 9 (2014).
- [7] M. B. Fuller, S. T. Skidmore, R. M. Bustamante, and P. C. Holzweiss, *The Review of Higher Education* **39**, 395 (2016).
- [8] C. Ngai, M. Pilgrim, D. Reinholz, J. Corbo, and G. Quan, *CBE life sciences education* **19**, ar15 (2020).
- [9] R. F. Woolson, *Wiley encyclopedia of clinical trials* pp. 1–3 (2007).
- [10] D. Sachmpazidi, C. Turpen, J. Petrella, R. P. Dalka, and F. N. Abdurrahman, *Phys. Rev. Phys. Educ. Res.* **20**, 010132 (2024), URL <https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.20.010132>.
- [11] J. C. Corbo, D. A. Craig, R. P. Dalka, and C. Turpen, in *Proceedings of the Physics Education Research Conference (PERC)* (2022), pp. 112–117.
- [12] D. Sachmpazidi, C. Turpen, and R. P. Dalka, in *Proceedings of the Physics Education Research Conference (PERC)* (2022), pp. 401–406.
- [13] F. N. Abdurrahman, C. Turpen, and D. Sachmpazidi, in *Proceedings of the Physics Education Research Conference (PERC)* (2022), pp. 24–29.