Examining reasons undergraduate women join physics

Maxwell Franklin and Eric Brewe
Drexel University, Philadelphia, Pennsylvania 19104, USA

Annette R. Ponnock
Yale University, New Haven, Connecticut 06520, USA

(Received 27 June 2022; accepted 3 November 2022; published 21 February 2023)

This study examines survey data from 2127 undergraduate women at the 2015 and 2019 American Physical Society Conference for Undergraduate Women in Physics (CUWiP) in order to classify what led them to study physics. We use expectancy-value and self-efficacy theory to create a coding framework based on different types of value and efficacy expectations in order to group responses. We find that the most common attractions are social persuasion, which is due to pressure or persuasion from people around the students, and intrinsic value, which is related to the inherent value of engaging in physics. Once the responses have been classified, we examine a follow-up survey to study whether the different motivational factors affect retention. We find that students who join physics because of the community are less likely to remain in physics after finishing their undergraduate studies. This is the first stage of a longer project to study which qualities correlate with retention in undergraduate women.

DOI: 10.1103/PhysRevPhysEducRes.19.010110

I. INTRODUCTION

There is a vast body of research around why women are driven out of physics. A hostile climate and negative stereotypes push women out of the field at each stage, resulting in a field that is overwhelmingly male [1–6]. This lack of gender diversity is harmful to the field, as it results in a lack of diversity of ideas as well as an unsupportive culture for women in physics. Though this work is very important, it is also crucial to study why women join physics in the first place, a question that has seen significantly less research. In this work, we examine what initially attracts women to physics. Unlike previous studies which view this through a lens of desired outcome [7], We do this by leveraging both expectancy-value theory and self-efficacy. Expectancy-value theory states that people are motivated to do a task by weighing the value of the task and their belief of success against the costs contained in that task. Self-efficacy refers to a person's belief in their ability to perform a task.

Previous work has shown that self-efficacy is a factor in choosing physics [8–10], but expectancy-value theory has only been used to explain persistence [11–14], not initial attraction. We expand previous literature by examining reasons women join physics through the lens of values as well as expectations of success. To this end, we analyzed survey responses collected at the 2015 and 2019 Conferences for Undergraduate Women in Physics (CUWiP). This research provides insight into reasons women join physics that could be used to attract more women to the field.

Additionally, we analyzed a follow-up survey given to the CUWiP attendees that tells us whether they stayed in physics through their undergraduate education and beyond. From this, we are attempting to make a predictive tool that finds which answers on the initial survey correlate with retention. This will allow us to give the survey to all undergraduate women and find which students need support in order to stay in the discipline. For this work, we found how likely students expressing each motivational code were to graduate with a degree in physics and remain in physics afterward.

We understand that gender is not a binary and that people who do not identify as women regularly attend CUWiP. However, very few attendees responded that they did not identify as women, and to avoid identifying anyone, we choose to follow the nomenclature of the conference and refer to attendees as women in this paper. That being said, it is possible that all people who are not cisgender men are represented in our data.

II. THEORY

We have chosen to use a framework for motivation based around expectancy-value theory, which states that

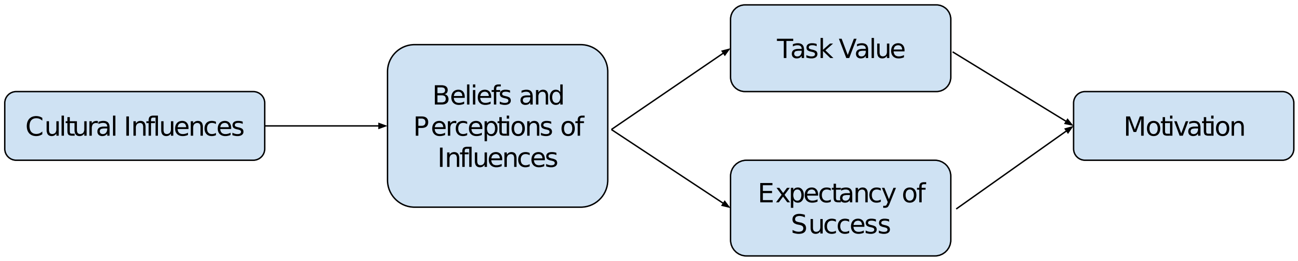

FIG. 1. The expectancy-value model of motivation. Our work focuses on the third stage, where expectancies of success, defined through Bandura's efficacy expectations, and task value, defined by Eccles, go together to create motivation.

motivation is based on two different factors: the value a person gives to a certain task and their expectancy for success. Both of these factors arise from cultural influences, which are then shaped by individual beliefs and perceptions [15,16]. These origins give expectancy-value theory both internal and external roots that interact to provide motivation [17]. This process is depicted in Fig. 1, which shows how the various factors combine to form motivation. In this work, we ignore the initial sources of motivation and look solely at the values and expectancies people hold.

A. Expectancy-value theory

Eccles proposed three different types of value when describing expectancy-value theory [15,18]. The first is attainment value, which occurs when the act of completing a task reinforces a person's identity. The second is intrinsic value, which means that a person finds the act of performing the task to be inherently valuable. The third is utility value, meaning that the task is useful for a person's future goals. In the survey responses that we collected, students showed each of these values in regard to studying and practicing physics.

These values are offset by costs, which are defined as anything that prevents a person from pursuing a task [15,19]. In the original proposal of expectancy-value theory, there were three types of cost. The first is task-effort cost, which reflects the difficulty of pursuing a task. The second cost is emotional cost, which happens when a person experiences negative feelings resulting from struggle or failure. The final cost in the original model is loss of valued alternatives cost, which is the external things given up in order to pursue a task. Further research added a fourth cost, outside effort cost [20,21]. This cost is the effort needed for other activities that distract from the task. Between values and costs, we fully cover the factor of value.

Previous work has used expectancy-value theory to study the difference in initial motivation between men and women [22]. This work found that men are overwhelmingly motivated to study biology because they find it to be interesting (relating to intrinsic value) and useful (utility value). On the other hand, they did not find any dominant trends for women. There has been no equivalence of this study for the physics community.

B. Self-efficacy theory

The second factor that determines motivation, expectancy for success, is best understood through self-efficacy theory. Self-efficacy is defined as a person's belief in their ability to perform a task, and it is mediated through efficacy expectations [23]. Efficacy expectations are experiences or feelings people have that shape how they view their own ability. When defining self-efficacy theory, Bandura posited four main sources of efficacy expectations [23]. The first is mastery experience, which is any experience in which a person succeeds at something related to the given task. The second expectation is vicarious experience, where a person sees someone they relate to succeed. The third is social persuasion, where other people give pressure or support to complete a task. The final efficacy expectation is physiological or emotional arousal, where performing the task makes the person feel good. From these four sources, people create an expectation for success that feeds into engagement. Past work has shown that women are more likely to express self-efficacy beliefs because of social persuasion and vicarious experience [24,25].

There is previous research both on what sources of efficacy expectations are most common in women [24] and on how expectancy-value theory explains gender differences in both non-academic [26,27] and academic settings [28,29]. These theories have been used extensively to study retention of women in academia [30], especially in science, technology, engineering, and mathematics (STEM) fields [11,12,31]. Our work combines both theories in order to fully explain the motivational factors that cause students to initially join physics and narrows the scope of the research specifically to undergraduate women in physics.

III. METHODOLOGY

Our data are composed of responses to two surveys. The first was given to attendees of the Conference for Undergraduate Women in Physics in 2015 and 2019. We used responses to a question that instructed students to

"Briefly describe events/experiences that led you to becoming a physics or physics-related major starting from the first time you considered physics as an option to the present (please include details of the factors that influenced you initially and along the way)." Between the two years, we had 2127 responses. The second survey was a follow-up survey designed to tell how students had progressed in their careers. The follow-up survey was sent to CUWiP participants four years after they attended the conference. Because of the recency of our data, all responses to the follow-up survey come from the 2015 cohort. Of the students who filled out the initial survey in 2015, 107 also completed the follow-up. We analyzed these data in three primary ways. First, we conducted a qualitative coding of all 2127 responses, described in Sec. III A.

A. Coding of open-ended survey responses

We began our coding scheme by assigning a distinct code to each type of value, cost, and efficacy expectation. As we began to code, the *intrinsic value* code proved to be too broad and subcodes emerged. Ultimately, our scheme is a mix of *a priori* and emergent coding, though it is mainly the former. This means though it was mainly defined beforehand, some codes developed organically as we went through the data. Three authors refined the coding scheme until interrater reliability passed 90%.

B. Coding responses

When coding responses to the initial survey, we created overarching categories based on the two motivation theories used in our framework: expectancy-value theory and self-efficacy theory. Each of these categories had a group of codes corresponding to the components of each theory. The codes in the expectancy-value category matched the types of value and cost put forth by Eccles and Flake [15,20], while the codes in the self-efficacy category matched Bandura's efficacy expectations [23].

In the *expectancy-value* category, we had four codes and numerous subcodes. The first code was *attainment value*, which was for responses that discussed a student's identity as a physicist or a scientist. The second was *utility value*, which was for responses about future goals as a physicist. This was a broad category, as it covered both career goals and less concrete objectives, such as making physics a more welcoming place for women. The third code was *intrinsic value*, which was for responses that spoke of the inherent value of doing physics. This was divided into *general intrinsic value* and categories for each specific field of physics. The final code was *cost*, which contained excerpts about what students gave up to do physics or what obstacles they encountered in studying physics.

The intrinsic value code is by far the broadest, so we added subcodes in order to make it more specific. In our initial round of coding, we added groups based on the specific discipline of physics that students found valuable. For example, there were many responses in which students talked about the inherent value of astrophysics. While there was a scattering of responses for other fields of physics, none besides astronomy were found in a significant amount of codes.

After our first round of coding, we felt that the intrinsic value code was still too general and we added three more subgroups based on what triggered the value. The first new subgroup is *event-triggered intrinsic value*, where a student found value in physics because of an event such as a class or science fair. The second subgroup is *media-triggered intrinsic value*, in which students valued physics because of persuasion from media sources. The third is *social intrinsic value*. This is distinct from other intrinsic value codes because it does not involve the student finding physics inherently valuable. Instead, this code refers to a response where a student finds value in being part of the physics community, for one reason or another. It is similar to both social persuasion and vicarious experience, but it differs in that it does not involve specific individuals influencing the student.

Our other most general code in the expectancy-value category is that of cost. This code is a good indicator of possible reasons women do not join or leave physics, and we split it into the subcategories defined by Eccles *et al.* and Flake [15,21]. The first is emotional cost, which describes the emotional or psychological toll that physics takes. This could be as simple as a lack of enjoyment in a physics class, but we often saw it tied to gender inequities in a physics department that create a hostile environment for women. The second cost is task-effort cost, which deals with the challenging nature of undergraduate physics. The third type of cost, the loss of valued alternatives, includes responses where students had to give up something, which is often a career or a family member's approval, in order to study physics. The fourth and rarest type of cost in our data is an outside effort cost, which describes external factors that take time away from doing physics such as family or health.

The 'self efficacy' group contained four categories, 'mastery experience', 'vicarious experience', 'social persuasion', and 'emotional and physiological arousal', which match the four sources of efficacy expectations. Codes in the first category, 'mastery experience', discussed previous experiences in which the student had done well in a physics or STEM setting. Codes in 'vicarious experience' contained experiences where a student saw others like them, such as a peer or relative, excel at physics. This code was later expanded to include any non-active influence from another person, such as a teacher creating a welcoming environment in a physics course. Codes in 'social persuasion' discussed active pressure or support from others, such as a teacher, relative, or friend, to join physics. Finally, 'emotional and physical arousal' contained excerpts where students detailed how performing physics made them feel good. A full guide to our coding scheme can be found in Table I.

TABLE I. Description of our coding framework.

Group	Category	Description	Example
Expectancy-value	Attainment value	Students pursue physics to further their identity as a scientist or as a physicist	"I've been the 'math and science girl' "
	Cost	Students are discouraged from pursuing physics	"I often question whether I'm in the right major because it's so hard"
	Intrinsic value	Students study physics for the value of physics or of science	"I was amazed by the explanatory and predictive power of physics"
	Media triggered intrinsic value	Students show intrinsic value because of a piece of media they consumed	I initially considered studying physics in 5th grade when I got a popular science book on quantum mechanics."
	Event triggered intrinsic value	Students show intrinsic value because of an event such as a camp, job, or research experience	"I first became fascinated with Physics and Astronomy as a young child when I went on a field trip to a planetarium"
	Intrinsic value(Social)	Students study physics because being part of the physics community is inherently valuable	"The people I had met in the physics department were generally more dynamic, the jokes were funnier, and there were more fun things to play with."
	Utility value	Students pursue physics to achieve a goal. This could be a career, success in another field, or to inspire others	"Since all engineering fields require a good physics background, I decided to pursue a physics degree"
Self-efficacy	Mastery experience	Students had an experience where they excelled at physics or a related field	"I realized that in my high school physics course I was doing better than other students"
	Physiological or emotional arousal	Students feel good while doing physics	"In high school, Honors Physics, when it was the only class I sat in and felt happy"
	Social persuasion	Students study physics because of pressure or support from others, often teachers, peers, or relatives	"The math and physics community at my university has always been a great source of support"
	Vicarious experience	Students saw someone they relate to excel at physics. This could be a fellow student, a relative, or any other role model	"My best friend in high school's older sister was a physics major at Wellesley College and I thought, if she can do it I can do it too!"

C. Analysis of participant persistence

After the initial responses had been fully coded, we matched them with the follow-up survey. We gathered our data from three questions on this survey. The first asked the respondents for their current educational status. The second asked them to list any degrees they had obtained, including degree type and field. The third asked them for their field of employment. Once we had cleaned the data, we were left with 107 follow-up responses. We sorted these responses into five categories. The first category referred to students who had graduated with a degree in physics and gone on to pursue physics in some field, either academia or industry. The second category contained students who had graduated with a degree in physics and had gone on to pursue a related STEM field. The third category referred to students who graduated with a degree in physics but did not continue in STEM. The fourth category had students who did not graduate with a degree in physics but had a degree in a related STEM field. The final category contained students who had not graduated with a degree in STEM. This included students who had obtained non-STEM degrees, students who had left college without obtaining a degree, and students still in their undergraduate experience.

Once the responses had been sorted, we found how prevalent each motivational code was in each response category and used these results to find the Bayesian posterior probability for each code. To do this, we begin with a prior model, which states how likely a random respondent is to fit into each category of retention. For each category, this is then multiplied by the likelihood that someone in the given category expressed the code in question. This is then divided by the probability that any given student, regardless of retention group, expressed the code. The final value tells us how likely someone who expresses a certain code is to be in each category of retention.

Though Bayesian analyses have not previously been used in this area of research to our knowledge, our reason for choosing them was twofold. First, Bayesian analyses focus on estimating the probabilities of an event; as a result, running multiple models does not inflate the likelihood of type one errors, or false positives, as it does in frequentist approaches. This is especially useful as we have a large number of codes, so there would be a lot of space for these errors otherwise. Second, Bayesian probability is especially good at dealing with small sample sizes. Though we had 107 responses to the follow up survey, only 13 respondents were in groups D or E, meaning they had not completed an undergraduate degree in physics. Furthermore, some of the rarer codes were only expressed a small number of times by respondents. These small sample sizes mean that other methods, such as statistical significance, would be meaningless and so Bayesian analysis is the correct choice.

We let P(A) denote the probability that a respondent answered a given category A, P(B) denote the probability that a respondent is in a retention category B, and L represent the likelihood function of A given B. The posterior probability that they are in retention category B given category A can be written as

$$\mathrm{P(B|A)} = \frac{\mathrm{P(B)L(B|A)}}{\mathrm{P(A)}}.$$

IV. RESULTS AND ANALYSIS

Here we describe the results of both the hand coding and the Bayesian analysis of persistence.

A. Motivational codes

Intrinsic value was by far the most prevalent code in the responses. General intrinsic value was found in 41.8% of responses, and event triggered intrinsic value was found in 21.4% of responses. Social persuasion, physiological or emotional arousal, and vicarious experience were the next most common codes, found in almost 30% of responses. The prevalence of vicarious experience and social persuasion are consistent with previous research [24]. The intrinsic value of astronomy was the only other code that was found in over 20% of responses. Exact percentages for each code, both by year and overall, can be found in Table II. Though we do not see any overwhelming differences between the two years, there are small shifts. Vicarious experience, general intrinsic value, and event triggered intrinsic value all saw small declines between 2015 and 2019, while the intrinsic value of astronomy had a small increase. The most drastic shift was that of the loss of valued alternatives subcategory of cost, which declined from being in 3.1% of responses to being in 1.1%.

One result to note is the rarity of cost. Emotional cost was found in less than 10% of responses, and no other type of cost was found in over 3% of responses. Although this is most likely due to our survey question, which asked for attractions to physics and was not designed to elicit costs, it is encouraging that so few women mentioned the barriers they faced to performing physics. It is also encouraging that the amount of people expressing costs shrank between 2015 and 2019, primarily in emotional cost and loss of valued alternatives.

TABLE II. Percentage of students who expressed each code. Self-efficacy codes are bolded, while expectancy-value codes are in normal text.

Year	2015	2019	Total
Intrinsic-general	44.99%	39.88%	41.82%
Social persuasion	**29.17%**	**30.44%**	**29.96%**
Physiological or emotional arousal	**29.54%**	**29.91%**	**29.77%**
Vicarious experience	**32.63%**	**23.87%**	**27.19%**
Event-triggered intrinsic value	25.09%	19.18%	21.43%
Intrinsic-astronomy	16.81%	22.21%	20.16%
Utility value	17.55%	15.94%	16.55%
Mastery experience	**15.57%**	**13.29%**	**14.16%**
Media-triggered intrinsic value	15.33%	13.22%	14.02%
Intrinsic value (Social)	10.88%	11.18%	11.06%
Cost-emotional	9.15%	6.57%	7.55%
Intrinsic-other subject	3.46%	3.10%	3.22%
Cost-task effort	2.60%	2.87%	2.77%
Cost-loss of valued alternatives	3.09%	1.06%	1.83%
Cost-outside effort	1.73%	1.74%	1.73%
Attainment value	1.98%	1.06%	1.41%

Our coding framework had two main sections: codes regarding the value students found in physics and codes regarding a student's self-efficacy towards performing physics. We did not find a clear discrepancy between the two overarching categories, as both had codes that were common and codes that were rare. From this, we conclude that there is not a bias in the stage of expectancy-value theory that influences women to join physics. As many women join based on their expectation of success as from the perceived value of the work.

B. Follow-up results

We found that the majority of students who took the follow up survey (about 49%) continued in physics after getting their degree. The next most populated group (28%) was composed of students who got a physics degree and continued in a different STEM field. The other groups only contained a handful of students.

Table III shows the posterior probability of an occurrence of a given code being in each category. If all the codes were evenly distributed, this probability should be the same as the overall proportion of codes. Because the two are

TABLE III. Bayesian posterior probability for each code and retention category. The top row is what proportion of students was in each category. Probabilities more than 5% above the overall proportion are italicized, while probabilities more than 5% below the overall proportion are bolded.

Category	Remain in physics	Graduate in physics, remain in STEM	Graduate in physics, leave STEM	Graduate in STEM	Don't graduate in STEM
Overall proportion	0.486	0.280	0.112	0.065	0.056
Social persuasion	*0.577*	**0.192**	0.115	0.038	0.076
Vicarious experience	0.5	0.281	0.062	*0.124*	0.031
Mastery experience	**0.368**	*0.368*	0.105	0.052	0.105
Physiological or emotional Arousal	0.459	0.270	0.108	0.054	*0.108*
Attainment value	0.5	0.249	*0.249*	**0**	**0**
Utility value	**0.421**	0.263	*0.263*	0.052	**0**
Intrinsic-general	0.531	0.285	0.102	0.081	**0**
Intrinsic value (astronomy)	**0.353**	0.294	*0.235*	0.058	0.059
Intrinsic value (other subject)	*0.571*	**0**	0.143	*0.142*	*0.143*
Intrinsic value (social)	**0.167**	*0.5*	*0.167*	0.083	0.083
Event-triggered intrinsic value	0.470	0.235	**0.059**	0.058	*0.176*
Media-triggered intrinsic value	0.526	0.237	0.131	0.078	0.026
Emotional cost	**0.375**	0.25	**0.25**	*0.124*	**0**
Cost (other)	*0.668*	*0.333*	**0**	**0**	**0**

different, there is some relationship between motivation and retention.

We can see that social persuasion is slightly overrepresented in students that remained in physics postgraduation and slightly underrepresented in students that remained in STEM. Meanwhile, vicarious experience is slightly overrepresented in students that graduated with a STEM degree. Mastery experience is very underrepresented in students remaining in physics and overrepresented in students remaining in STEM. Utility value is overrepresented in students graduating in physics but leaving STEM. The intrinsic value of astronomy is underrepresented in students that remained in physics and overrepresented in students that graduated in physics but left STEM. The intrinsic social value is very underrepresented in students remaining in physics and very overrepresented in students remaining in STEM. The categories of cost, intrinsic value of a nonastronomy subject, and attainment value were not represented enough (< 10 instances) to give clear insight into their distribution.

V. DISCUSSION

A. Motivational codes

Our findings on the most common motivators of women offer paths for how to attract women to physics. As intrinsic value is the most common reason that women join physics, one possible way to attract more women to physics is to impart this value more widely. Our data show that the most likely trigger (outside of class or family) of intrinsic value is events, such as camps or research experiences. If these can be provided to more women at a young age, intrinsic value might be more widespread, encouraging more women to join physics. Another prevalent trigger was physics-related media, which can be leveraged in two ways. The first is by increasing the accessibility to physics media, such as by showing it in classes. The second is to diversify the types of media available. Almost all of the media figures referenced by the respondents were men, such as Michio Kaku or Stephen Hawking, but several women said that they were specifically brought in by seeing other women in physics. If books or videos created by women were more widely disseminated, it may increase the number of women in physics.

The prevalence of codes that relate to other people, especially vicarious experience and social persuasion, show the need to be intentional about fostering relationships with women in physics. The relative scarcity of the intrinsic value of physics shows that women who are influenced to join physics are mostly compelled to do so by individual relationships, rather than by the community as a whole. With this in mind, educators should try to not only foster a supportive environment but also to develop personal connections with their students. As vicarious experience and social persuasion tend to be more prevalent among women than men [32], this would disproportionately influence women to join physics, evening the gender ratio of men to women and in turn creating a more supportive environment for women.

B. Changes in cost over time

Though our survey question means that the rarity of cost in our results is not necessarily meaningful, the change in costs over time is still useful information. Because the question did not change between 2015 and 2019, it should have elicited costs just as much in each year. This leads to a very encouraging result. Task effort cost went up very slightly, but emotional cost fell by almost one-third.

The most drastic shift between 2015 and 2019 was in the *other cost* category, which includes outside effort and loss of valued alternatives. While outside effort cost stayed remarkably constant between the years, going from being found in 1.73% of responses to 1.74%, the loss of valued alternatives cost plummeted 3.09% to 1.06%, by far the biggest change of any code. This manifested as many women in 2015 saying that they were not aware of viable career paths in physics, while almost no women expressed that belief in 2019. Though we cannot ascertain any reasons for this from the data, the four year period between surveys saw an increase in outreach from organizations such as APS. This includes the STEP-UP program, which discusses career paths in physics [33], and CUWiP panels discussing non-academic career paths in physics [34]. As the two different years sample an almost identical population, this is an encouraging result.

C. Retention

We saw clear trends in motivational categories expressed by women who continued in physics. It should be noted, however, that a small sample size hampers our ability to see trends among students who do not finish their undergraduate work with a physics degree. Our biggest finding from this work is that women expressing that they joined physics because of the community are much less likely to continue in physics after their undergraduate work. Despite the fact that almost half of our respondents remained in physics after graduating, only 16.7% of women who expressed the intrinsic value of community persisted. The implications of this are twofold. The first is women who show this motivational factor need to be given more support to stay in the field. The second is that the welcoming community that attracts women to physics is less present in later stages of physics; work should be done to make communities such as physics graduate school more collaborative and welcoming.

The other categories that heralded a woman leaving the field of physics were mastery experience and the intrinsic value of astronomy. Though these were less drastic, both had a retention rate ten percentage points lower than the average. In the case of mastery experience, the causes may be similar to social intrinsic value, as women are recognized less for success at higher levels of physics. To fix this, instructors need to communicate when women do well in physics or related disciplines. Regardless of reasons why these codes forebode a lack of retention, more support to stay in physics should be given to women who express these motivational codes.

VI. FUTURE WORK

A. Retention

We plan to continue our focus on the retention of women in physics. The question about motivational factors used in this research was part of a larger survey that included both demographic information and Likert scale questions about physics attitudes. We will study how the responses to each of these correlate with retention to obtain a full picture of the qualities of women that are retained in physics. Using this and machine learning, we will create a predictive tool for retention in order to see which undergraduate women might need support to remain in the department.

One limitation of our study is that the women who are likely to go to CUWIP tend to be more engaged in physics than the average undergraduate woman. This is clear in our data, as of the 107 students who responded to the follow-up survey, 94 of them had completed their degree in physics four years after attending the conference. Because of this, more research needs to be conducted in every aspect of our study. It is possible that most women in physics do not express the same motivational codes as the women who attended CUWIP, or that they express codes using different language so that our predictive tool would no longer be valid. Furthermore, the heavy trend towards graduating in our sample means that we have very little data on students who fail to graduate in physics at all. In order to obtain a more accurate picture of which motivational codes correlate with retention, we would need a larger sample size. We hope to conduct this research going forward

B. PYTHON tool

As coding all 2127 responses took us several days to complete, the process would be time prohibitive for others who wish to conduct this analysis. Because of this, we have developed a predictive tool using PYTHON that takes a written response as input and returns a vector describing the proportion of each motivational code present. This work was done concurrently with the work presented in this paper, and at the moment, our program shows a 65% accuracy when compared with the results of our hand coding. As we expand to further data sets, the larger amount of training data will allow us to obtain a higher accuracy. This tool will be necessary for people who want to repeat this study with their own students and do not have the background in motivation needed for hand coding.

C. Expanding to other data

Though we do not have the necessary data, our research raises the question of whether men would yield similar results to women. Previous research has found that men are more likely to express mastery experiences, while women are more likely to be motivated by vicarious experiences and social persuasion [32]. Though that finding was not specific to physics, it is likely that we would see a similar result. Another possible difference is in the specific words used when expressing each code. Our predictive tool, which is based on the most common words used by women, might not be applicable to men if they use different language to describe their experiences.

VII. CONCLUSIONS

In this study, we find that by far the most common reason women join physics is intrinsic value, meaning that physics is inherently worth studying. This has many implications for attracting women to physics; teachers should stress how physics connects to other subjects and explains the world around us. Furthermore, we see specific events, such as science fairs and camps, that foster this intrinsic value. We also confirm previous research findings that women are strongly influenced by their relationships with others [24], shown in the prevalence of social persuasion and vicarious experience codes. As we examine these survey responses going forward, we will use our PYTHON tool in order to quickly categorize the data.

We also find that most categories of motivation are not strongly linked to retention. The most prominent result from our data is that students who join for the intrinsic value of the physics community are much less likely to remain in physics after finishing their undergraduate work. Other codes such as joining physics because of a mastery experience, joining because of the intrinsic value of astronomy, or suffering an emotional cost are also less prevalent in students that remained in the discipline. Going forward, we will add the other questions on the CUWiP survey to this analysis in order to create a full picture of the qualities that enable women to keep pursuing physics through their undergraduate and professional careers.

ACKNOWLEDGMENTS

This work is supported by the National Science Foundation (PHY 2011766) and the NSF Graduate Research Fellowship (Fellow No. 2021314028).

[1] B. Bloodhart, M. M. Balgopal, A M. A. Casper, L. B. Sample McMeeking, and E. V. Fischer, Outperforming yet undervalued: Undergraduate women in STEM, PLoS One **15**, (2020).

[2] Jennifer Blue, Adrienne L. Traxler, and Ximena C. Cid, Gender matters, Phys. Today **71**, 3, 40 (2018).

[3] Anne Marie Porter and Rachel Ivie, *Women in physics and astronomy* (American Institute of Physics Statistical Research Center, College Park, MD, 2019).

[4] Ramin Skibba, Women in physics face big hurdles—still, Nature(London) (2016).

[5] Danny Doucette and Chandralekha Singh, Why are there so few women in physics? Reflections on the experiences of two women, Phys. Teach. **58**, 297 (2020).

[6] Jessica Ellis, B. K. Fosdick, and C. Rasmussen, Women 1.5 times more likely to leave STEM pipeline after calculus compared to men: Lack of mathematical confidence a potential culprit, PLoS One **11**, (2016).

[7] Linda J. Sax, Kathleen J. Lehman, Ramón S. Barthelemy, and Gloria Lim, Women in physics: A comparison to science, technology, engineering, and math education over four decades, Phys. Rev. Phys. Educ. Res. **12**, 020108 (2016).

[8] Tesfaye Semela,Who is joining physics and why? Factors influencing the choice of physics among Ethiopian university students, Int. J. Environ. Sci. Educ. **5**, 319 (2010).

[9] Stephen R. Porter and Paul D. Umbach, College major choice: An analysis of person–environment fit, Res. High. Educ. **47**, 429 (2006).

[10] Zheng Zhu, Learning content, physics self-efficacy, and female students' physics course-taking, Int. Educ. J. **8**, 204 (2007).

[11] Ghada Nehmeh and Angela Kelly, Women physicists and sociocognitive considerations in career choice and persistence, J. Women Minorities Sci. Engin. **24**, 95 (2018).

[12] Joseph Appianing, Using the expectancy-value theory to understand why women persist or leave collegiate stem programs: A mixed methods study, Theses and Dissertations, 2096 (2017).

[13] Ladewig, Antonia, Olaf Köller, and Knut Neumann, Persisting in physics and the physics olympiad—impact of gender identification and sense of belonging on expectancy-value outcomes, Eur. J. Psychol. Educ. (2022).

[14] Cynthia A. McGrath, K. Gipson, O. Pierrakos, R. Nagel, J. Pappas, and M. Peterson, An evaluation of freshman engineering persistence using expectancy-value theory, in *Proceedings of the 2013 IEEE Frontiers in Education Conference* (FIE, Piscataway, NJ, 2013).

[15] Jacquelynne Eccles, T. F. Adler, R. Futterman, S. B. Goff, C. M. Kaczala, J. Meece, and C. Midgley, Expectancies, values and academic behaviors, in *Achievement and Achievement Motives*, edited by J. T. Spence (W. H. Freeman, San Francisco, 1983).

[16] Allan Wigfield and Jacquelynne S. Eccles, Expectancy-value theory of achievement motivation, Contemp. Educ. Psychol. **25**, 68 (2000).

[17] Allan Wigfield, Expectancy-value theory of achievement motivation: A developmental perspective, Educ. Psychol. Rev. **6**, 49 (1994).

[18] Allan Wigfield and Jacquelynne S. Eccles, The development of achievement task values: A theoretical analysis, Developmental Rev. **12**, 265 (1992).

[19] Kenneth E. Barron and Chris S. Hulleman, Expectancy-value-cost model of motivation, in *International Encyclopedia of the Social and Behavioral Sciences* (Elsevier, 2015).

[20] Jessica Kay Flake, Measuring cost: The forgotten component of expectancy value theory, master's thesis, James Madison University, Harrisonburg, VA, 2010–2019. 208 (2012).

[21] Jessica Kay Flake, Kenneth E. Barron, Christopher Hulleman, Betsy D. McCoach, and Megan E. Welsh, Measuring cost: The forgotten component of expectancy-value theory, Educ. Psychology **41**, 232 (2015).

[22] Ellen S. Sullins, Delia Hernandez, Carol Fuller, and Jay Shiro Tashiro, Predicting who will major in a science discipline: Expectancy-value theory as part of an ecological model for studying academic communities, J. Res. Sci. Teach. **32**, 99 (1995).

[23] Albert Bandura, Self-efficacy: Toward a unifying theory of behavioral change, Psychol. Rev. **84**, 191 (1977).

[24] Amy L. Zeldin and Frank Pajares, Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers, Am. Educ. Res. J. **37**, 215 (2000).

[25] Vashti Sawtelle, Eric Brewe, and Laird H. Kramer, Exploring the relationship between self-efficacy and retention in introductory physics, J. Res. Scientific Teach. **49**, 1096 (2012).

[26] J. S. Eccles and R. D. Harold, Gender differences in sport involvement: Applying the Eccles' expectancy-value model, J. Appl. Sport Psychol. **3**, 7 (1991).

[27] Zan Gao and Ping Xiang, College students' motivation toward weight training: An application of expectancy-value model, J. Teaching Phys. Educ. **27**, 399 (2008).

[28] Weihua Fan, Social influences, school motivation and gender differences: An application of the expectancy-value theory, Educ. Psychol. **31**, 157 (2011).

[29] Oliver Dickhäuser and Joachim Stiensmeier-Pelster, Gender differences in the choice of computer courses: Applying an expectancy-value model, Social Psychol. Educ. **6**, 173 (2003).

[30] Theresa Schnettler, Julia Bobe, Anne Scheunemann, Stefan Fries, and Carola Grunschel, Is it still worth it? Applying expectancy-value theory to investigate the intra-individual motivational process of forming intentions to drop out from university, Motivation and emotion **44**, 491 (2020).

[31] Ming-Te Wang and Jessica Degol, Motivational pathways to STEM career choices: using expectancy-value perspective to understand individual and gender differences in STEM fields, Developmental Rev. **33**, 304 (2013).

[32] Amy Zeldin, Sources and effects of the self-efficacy beliefs of men with careers in mathematics, science, and technology, Doctoral Dissertation, Educational Studies, Emory University, 2000.

[33] STEP-UP American Physical Society (2020).

[34] Crystal Bailey (personal communication).