



PAPER • OPEN ACCESS

## The relationship between gender and academic performance in undergraduate physics students: the role of physics identity, perceived recognition, and self-efficacy

To cite this article: Ewan Bottomley *et al* 2023 *Eur. J. Phys.* **44** 025701

View the [article online](#) for updates and enhancements.

### You may also like

- [Model Development in Predicting Academic Performance of Students Based on Self-Efficacy Using K-Means Clustering](#)  
Mohammad Zahid Mohammad Sabri, Nazatul Aini Abd Majid and Siti Aishah Hanawi
- [The important of self-efficacy and self-regulation in learning: How should a student be?](#)  
U Toharudin, A Rahmat and I S Kurniawan
- [Gender differences in test anxiety and self-efficacy: why instructors should emphasize low-stakes formative assessments in physics courses](#)  
Alysa Malespina and Chandralekha Singh

# The relationship between gender and academic performance in undergraduate physics students: the role of physics identity, perceived recognition, and self-efficacy

Ewan Bottomley<sup>1,2,\*</sup> , Antje Kohnle<sup>1</sup>, Kenneth I Mavor<sup>2</sup>,  
Paula J Miles<sup>2</sup> and Vivienne Wild<sup>1</sup>

<sup>1</sup> School of Physics and Astronomy, University of St. Andrews, United Kingdom

<sup>2</sup> School of Psychology and Neuroscience, University of St. Andrews, United Kingdom

E-mail: [eb301@st-andrews.ac.uk](mailto:eb301@st-andrews.ac.uk)

Received 23 March 2022, revised 27 October 2022

Accepted for publication 11 November 2022

Published 26 December 2022



CrossMark

Keywords: physics identity, gender, self-efficacy, perceived recognition as a physicist, academic performance

(Some figures may appear in colour only in the online journal)

## Abstract

Studies focusing on physics undergraduate students have found that women tend not to identify as strongly with physics, compared to men. Recent research has examined potential factors that influence the experience of women in physics. Several of these factors, such as students' beliefs in their ability to complete physics-based tasks (i.e., self-efficacy) and students' belief that others perceive them as a physicist (i.e., perceived recognition), have been associated with physics identity in the context of introductory university physics courses in the United States (US). The current study extends this previous work, surveying students at all levels of the undergraduate degree at a research-intensive university in the UK. Students were asked about their physics identity, physics self-efficacy, and the extent to which they

\* Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

believed others perceived them as physicists. The survey responses were then matched with students' grades. Using matched responses from the start and end of an academic year from 169 students (110 men, 59 women), two analyses were performed. The first analysis found that average scores for women for physics identity and self-efficacy were lower than for men both at the start and end of the academic year. The second analysis found that after controlling for the start-of-year scores in physics identity, self-efficacy, and perceived recognition, students' mid-year grades significantly predicted variance in their end-of-year scores for self-efficacy, perceived recognition, and (possibly also) physics identity. This study also found that the gap in perceived recognition between men and women increased over the academic year. The results contribute to understanding potential barriers for women in physics and have implications for instruction in terms of promoting students' physics identity, self-efficacy, and perceived recognition.

## Introduction

The lived experience of women in the undergraduate physics classroom can be very different from the experience of men. Societal expectations and stereotypes can facilitate negative experiences for women e.g. [1]. Moreover, the negative experiences that women face can undermine their identity as scientists [2]. At the university level, students' sense of identification with their discipline of study may be important to their academic outcomes, with discipline identity predicting students' willingness to continue to study their chosen subject in the future e.g. [3] and the approaches they take to their learning [4]. This is problematic as recent research at the introductory undergraduate level in the US has suggested that women do not identify with physics as a discipline to the same extent as men, in part, due to women having lower self-efficacy (i.e. being less confident in their abilities to perform physics tasks) and having lower perceived recognition (i.e. not feeling as recognized as a physicist) [5, 6]. However, how physics identity relates to academic performance at the university level has been sparsely studied, particularly outside of the US education system. As such, the current study aims to investigate the gender differences in physics identity, self-efficacy, and perceived recognition as a physicist, across all levels of the undergraduate degree at a UK institution. We also aim to examine whether academic performance can predict these factors.

### *Gender framework*

In the text that follows we discuss the concept of gender in relation to comparisons between men and women. We have focused on men and women as they were the only genders for which we could collect a big enough sample to adequately power our analyses, and this reflects the distinction of long-standing concern in terms of women's participation in physics. It is not possible to draw conclusions from our data from the small numbers of non-binary identified participants. The evolving experience of those students identifying as non-binary is an interesting question for future research but is beyond the scope of the current paper.

### *Gender and physics identity*

In their influential paper, Hazari, Sonnert, Sadler, and Shanahan [7] examined physics identity in a sample of students in their first year of college. The students were asked to reflect on their high school experiences and complete a survey measuring their identification with physics as a discipline. Hazari and colleagues found that the students that reported greater physics identity were more likely to indicate that they wanted to follow a physics-related career path

in the future. However, when examining physics identity by gender, Hazari and colleagues showed that women reported significantly lower physics identity than men. Therefore, the authors concluded that women would be less likely than men to follow a physics career pathway as they do not identify with physics as a discipline to the same extent as men. This suggests that the under-representation of women in physics may be a product of women's relatively low physics identity.

Following this, physics education researchers began to examine the factors that form physics identity. Kalender and colleagues [6] highlighted several factors which may contribute to gendered patterns in the formation of physics identity in introductory university physics courses, in particular, physics self-efficacy and perceived recognition as a physicist. Specifically, two of the many factors that they suggest are contributing to women's lower physics identity are physics self-efficacy (a students' belief in their abilities to complete physics-based tasks) and perceived recognition as a physicist. With regards to self-efficacy, there is a large body of literature that suggests that women are generally not as confident in their abilities in physics e.g. [8–10]. For instance, Marshman and colleagues [11] found that women who were achieving an A-level grade in an introductory college physics course had similar levels of self-efficacy as men who were achieving a C-level grade. This is problematic as physics self-efficacy has been shown to predict levels of physics identity [6]. Therefore, gender disparity in self-efficacy may be contributing to the gender difference in physics identity.

The extent to which university students feel recognized as a physicist by others is also important for the formation of physics identity [5, 6, 12, 13]. Kalender and colleagues [5] showed that students' beliefs that their friends, family, and instructors saw them as a physicist was associated with greater interest in physics, competency beliefs, and physics identity. In their study, they found that men reported greater perceived recognition as a physicist than women. As such, Kalender and colleagues concluded that perceived recognition as a physicist, much like self-efficacy, potentially contributes to the gender differences in physics identity.

#### *Academic performance and students' perceptions*

University students' identification with their discipline of study has been related to their academic performance e.g. [3, 14, 15]. Seyranian and colleagues [16] examined this relationship in an introductory college physics course. They asked students to complete a survey at the beginning and at the end of the course measuring students' sense of belonging and physics identity. The results showed that men tended to report greater belonging and physical identity than women. Students who reported more physics identity tended to earn better grades. Moreover, the students that performed better in their exams generally reported more physics identity at the end of the academic year. Thus, Seyranian and colleagues concluded that there was a bidirectional relationship between physics identity and academic performance.

There is a large body of work examining the relationship between self-efficacy and academic performance e.g. [17, 18]. In particular, Kalender and colleagues [6] hypothesize that the relationship between self-efficacy and academic performance may be similarly bidirectional, in what they describe as a 'feedback loop'. They argue that students who are more confident in their abilities tend to get better grades, and better grades will encourage the students to be more confident in their abilities. However, Kalender and colleagues [6] did not directly examine this relationship. Such a feedback loop would be problematic if it exists as gender disparity in physics identity and self-efficacy may result in lower academic

performance for women, which may, in turn, negatively impact women's sense of self-efficacy and physics identity in the future.

Despite the conjecture around potential feedback loops, recent work by Whitcomb and Singh [19] suggests that physics students' academic performance in the introductory years of study is not predictive of the students' academic performance at more advanced levels of the undergraduate degree. They used structural equation modelling to demonstrate that gender differences do emerge in the grades of introductory level physics students, with men tending to perform slightly better, but that the introductory physics grades a student received did not predict the grades the student received later in their academic trajectory. The authors posit that women's desire to continue with physics may be impacted by their relatively lower grades at an early stage. What remains unclear is whether a physics students' academic performance predicts their perceptions in physics (e.g. their physics identity, self-efficacy, perceived recognition), which may contribute to women not continuing their physics career. Therefore, this study aims to examine the relationship between grades and student's motivational factors in undergraduate physics.

### *Study overview*

The study presented here aimed to build on the work of Kalender and colleagues [5, 6], Whitcomb and Singh [19], and Seyranian and colleagues [15] by investigating gender differences in physics identity, self-efficacy, and perceived recognition and the association these factors may have with academic performance across all levels of the undergraduate physics degree in a UK context. Our study aimed to address the following two research questions:

- (1) Are there gender differences in self-efficacy, perceived recognition and physics identity, and do they persist across the academic year?
- (2) Does academic performance predict self-efficacy, perceived recognition and physics identity in the following semester?

We conducted two analyses, each addressing one of these research questions. Whilst previous studies have predominantly investigated students' experience of physics in the US at the high school or introductory university level where only a minority of the students in these courses intended to pursue physics, the study presented here focuses on physics majors in the UK from the introductory to the advanced undergraduate level. Due to the longitudinal nature of our study, we were also able to investigate whether academic performance predicted self-efficacy, perceived recognition and physics identity (RQ2) in the following semester.

## **Study methodology**

### *Participants and data collection*

Participants were undergraduate students with degree intentions in physics (including joint degrees, such as mathematics and physics) at a small, selective, research-intensive university in the UK. Participants were recruited across all levels of the degree program, including the integrated Masters level. In what follows, we use 'level' and 'level of study' to denote the progression through the degree program, ranging from level 1 (introductory) to level 5 (integrated Masters); we use 'year' to denote the academic year of data collection, and we use 'timepoint' to denote the time of the individual rounds of data collection. We also used 'semester 1 and 2' to distinguish whether the surveys were collected at the start (semester 1) or at the end (semester 2) of an academic year. Participants were asked to complete surveys at

**Table 1.** A breakdown of the data collection timepoints. In brackets we indicate the number of surveys used in the matched analysis presented in this paper, as well as the percentage of responses from women in this analysis.

Timepoint	Levels surveyed <sup>a</sup>	Number of surveys completed (used in the matched analysis)	Percent of responses from women for completed surveys (surveys used in matched analysis)
1	1–5	309 (116)	34.0% (33.6%)
2	1–5	116 (116)	33.6% (33.6%)
3	3	82 (53)	31.7% (37.8%)
4	3	71 (53)	33.8% (37.8%)
5	3 and 4	139 (0)	30.2% (—)

<sup>a</sup> Levels surveyed refers to the progression through the degree program, ranging from level 1 (introductory) to level 5 (integrated masters).

one or more timepoints over three academic years. Students completed the surveys in class and were informed that their participation was voluntary. Excepting one timepoint where the surveys were completed online, the surveys were completed on paper. In total, data was collected at five different timepoints: the start of three academic years, and the end of two academic years.

A total of 449 students completed the survey at one or more of the timepoints. Fourteen completed surveys were removed as the survey did not include an identifier to match grades and gender information. We only wanted to focus on physics majors, so all participants for whom physics was not their primary or joint degree were removed ( $n = 42$ ). This left a remaining sample of 393 participants, with 717 completed surveys in total. A breakdown of surveys completed by timepoint and gender is shown in table 1. Excepting for timepoint 2 (where data was collected online), participation was high, with  $\sim 60\%$  to  $\sim 100\%$  of physics majors in each level of study completing the survey. For timepoint 2,  $\sim 15\%$  to  $\sim 50\%$  of physics majors in each level completed the study. While students at all levels were surveyed at timepoints 1 and 2, only students at level 3 (and in one case also level 4) were surveyed at timepoints 3–5. This was on the one hand due to practical limitations in terms of class time needing to be available for survey completion. On the other hand, level 3 is pivotal in terms of students' trajectory towards a BSc or integrated Masters degree, as the two pathways diverge after this level. Thus, this was also a level of particular interest.

For the study presented here, we focused on a subset of the data from timepoints 1–4, namely matched students with data from both the start and end of a single academic year (i.e. timepoints 1 and 2, or timepoints 3 and 4, see section 'Analysis Overview'). This included 169 students (110 men, 59 women, two surveys per student). As this may plausibly be a biased sample of students positively inclined to complete both surveys, we precede each analysis with results from the largest sample at timepoint 1. This timepoint surveyed ca. 80% of the total student population of physics majors.

### Survey measures

In what follows, where we use a related set of questions to measure a construct, we estimate internal reliability using Cronbach's alpha [20]. This allows us to gain an understanding of the internal consistency of the items we have used. A Cronbach's alpha typically ranges from 0 (no internal consistency) to 1 (complete internal consistency). The mathematical formula for the calculation of this statistic is:

$$\alpha = N * C / (V + (N - 1) * C),$$

where  $N$  = the number of items;  $C$  = the mean covariance between the items; and  $V$  = the mean item variance. The measure assumes that the scale in consideration is unidimensional and therefore, the larger the value, the more the variance in the items that are shared with other items and on average shows consistency in the thing being measured. The lower acceptable bound for internal reliability tends to be considered around .70 (see [21]). In the sections below we quote Cronbach's alpha for timepoint 1 only as a measure of internal reliability of the dataset, as this timepoint comprised the largest number of participants and therefore gives the best estimate of the reliability of the dataset.

All survey items were measured on a seven-point Likert Scale (1 = not at all/strongly disagree, 7 = very much so/strongly agree), excluding the self-efficacy measure at timepoints 1 and 2 which was measured on a five-point Likert Scale. A transformation was used on these timepoints such that we could compare self-efficacy across timepoints.

#### *Physics identity*

To measure physics identity, we used a single item from Hazari and colleagues' [7] Physics Identity Survey. This item reads: '*do you see yourself as a physics person?*'. This item has been used in several studies e.g. [5, 6], and was found to be a good proxy for overall physics identity e.g. [22].

#### *Perceived recognition as a physicist from others*

Perceived recognition as a physicist from others was also measured using the Physics Identity Survey [7]. This measure consisted of three items: '*do your physics teachers/instructors see you as a physics person?*'; '*do your parents/relatives/friends see you as a physics person?*'; and '*do your friends seek your advice/input in physics-related problems/discussions?*'. The Cronbach's alpha for the timepoint 1 dataset was .717.

#### *Self-efficacy*

To measure self-efficacy, we used the Global Self-Efficacy measure of the Physics Self-Efficacy Questionnaire [23]. This measure included four items (Cronbach's alpha timepoint 1 = .744): '*I will remain calm in my physics exam because I know I will have the knowledge to solve the problems?*'; '*I generally manage to solve difficult physics problems if I try hard enough?*'; '*I know I can stick to my aims and accomplish my goals in physics?*'; and '*I know I can pass the physics exam if I put in enough work during the semester?*'.

#### *Academic performance*

Students' academic performance was measured using their physics credit-weighted mean grade data for each semester. Grade information and gender was added to the surveys prior to anonymizing the data for analysis. Given that students were surveyed across all levels of the degree programme from the introductory to the Masters level, we used grades rather than standardized diagnostic instruments (such as the Force Concept Inventory) to measure academic performance. While individual courses may vary slightly in their mean grade, and there is variation in which courses students take at the more advanced levels, grades in the UK are measured on a common scale that takes UK degree classifications into account. Thus, mean grades are similar across the courses and levels. Please note that we use the terms academic performance and grades interchangeably throughout.

For the study presented here the measure of academic performance was the credit-weighted mean for the semester; therefore, the measure was calculated at the end of the semester. For RQ1 we used the credit-weighted mean score for the first (mid-year) and second (end-of-year) semesters; however, for RQ2 we used only the mid-year credit-weighted mean.

### *Analysis overview*

To examine the gender differences relating to RQ1 and the associations between grades and students' perceptions in RQ2, we used a sample that consisted of the students who responded to the survey at both the start and end of a single academic year. This allowed us to examine gender differences, differences in perceptions over time, and the effect of mid-year grades on perceptions. In the cases in which we had missing data, we used listwise deletion. This means that we only used the data of the participants that responded to all the survey items. A total of 169 students (110 men, 59 women) completed the survey at both the beginning and end of an academic year. Ten of these participants responded at the start/end of two academic years, so their second response was removed in order to ensure independence of measures.

To address RQ1 on gender differences between the constructs, we ran a series of 2 (gender: men versus women)  $\times$  2 (semester: semester 1 versus semester 2) ANOVAs (analysis of variance). This test examines the variance amongst and between groups to assess if there are significant differences between the means e.g. [24].

To examine whether academic performance is associated with physics identity, perceived recognition, or self-efficacy in the following semester (RQ2), we used a series of multiple regressions. This allowed us to see how much variance in the semester 2 scores is explained by the mid-year academic performance of students. The first step is to assess whether a student's semester 1 perception of, for example, self-efficacy predicts their perceptions of self-efficacy in the following semester. This is achieved by fitting a linear regression line. The regression equation for this model is:

$$Y' = M + \beta_1 * X_1,$$

where  $Y'$  = the predicted outcome (in our example case, predicted self-efficacy at semester 2);  $M$  = the  $Y$  intercept (the value of  $Y$  when everything else is considered 0);  $\beta_1$  = the regression coefficient associated with (in the case of our example) self-efficacy;  $X_1$  = the measured self-efficacy score at semester 1.

The next step is to add academic performance as an additional variable. This regression equation reads:

$$Y' = M + \beta_1 * X_1 + \beta_2 * X_2,$$

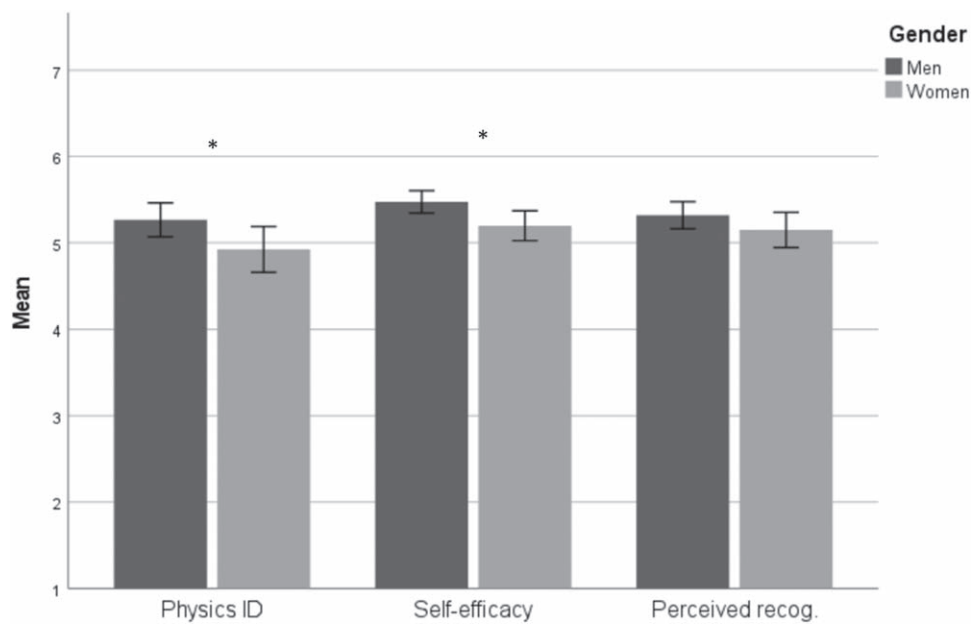
where  $\beta_2$  = the regression coefficient for the variable academic performance; and  $X_2$  is the observed mid-year academic performance. We can then assess whether the second model, with academic performance, is a better fit for our data than the first model without academic performance. If so, we can conclude that academic performance does predict the semester 2 perception (e.g. of self-efficacy), beyond that of the semester 1 perception.

## **Results**

### *RQ1: gender differences*

*Timepoint 1 dataset.* The matched dataset of 169 students (110 men, 59 women), consisting of students that responded both at the start and end of an academic session, is only a subset of the full dataset (see table 1), Therefore, we first checked global trends in gender differences





**Figure 1.** Mean scores in physics identity, perceived recognition, and self-efficacy for men and women using the full timepoint 1 dataset ( $n = 309$ , 204 men, 105 women). 'ID' represents physics identity and 'Perceived recog.' represents perceived recognition as a physicist. Error bars are 95% confidence intervals on the mean. \* Represents significant differences.

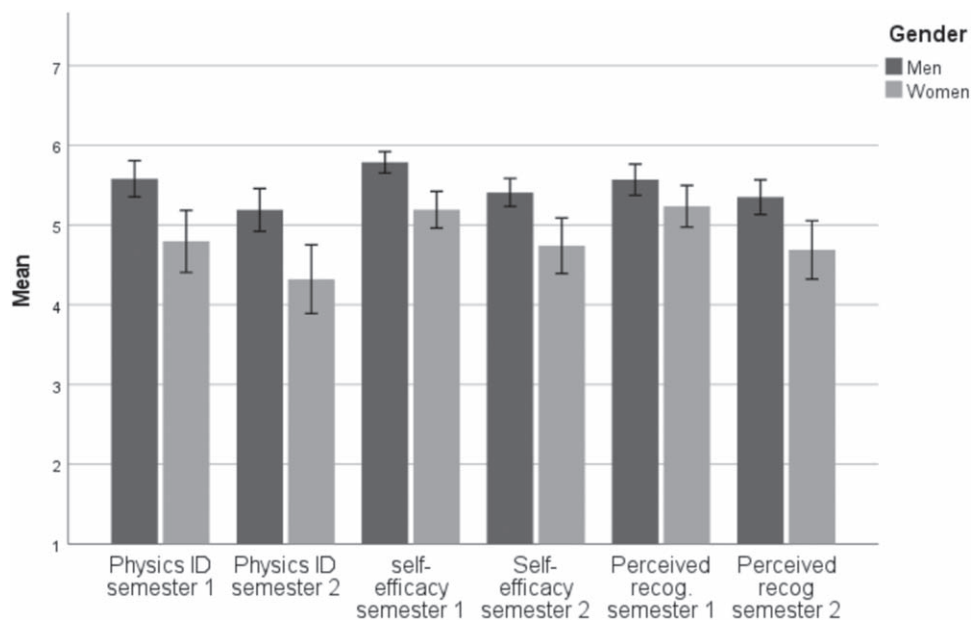
**Table 2.** The results of independent sample Mann–Whitney U tests from the first timepoint ( $n = 309$ , 204 men, 105 women). The  $p$ -value, and the means for men and women are reported. Significant results are highlighted in bold. Scores of perceptions were on a 7-point Likert Scale. At timepoint 1, self-efficacy was originally measured on a 5-point scale but has been adjusted to a 7-point scale for comparison.

Variable	$p$ -statistic	Mean men	Mean women
ID	<b>.016*</b>	5.27	4.92
Self-efficacy	<b>.010*</b>	5.48	5.20
Perceived recognition	.130	5.32	5.17

\* $p < .05$ .

\*\* $p < .01$ .

using the full sample from timepoint 1 (309 students, 204 men, 105 women). This timepoint surveyed ca. 80% of the total student population of physics majors. To examine whether there were gender differences in the timepoint 1 sample, we ran Mann–Whitney U tests (see figure 1 and table 2). Mann–Whitney U tests rather than t-tests were used as the data in this sample was not normally distributed. The results demonstrated that there were significant gender differences in physics identity and self-efficacy, with men reporting more physics identity and self-efficacy than women. Men also received slightly higher grades than women. No significant differences emerged between men and women with regards to perceived recognition.



**Figure 2.** Mean scores in physics identity, perceived recognition, and self-efficacy for men ( $n = 110$ ) and women ( $n = 59$ ) in the matched sample, at the start (semester 1) and end (semester 2) of an academic year. 'ID' represents physics identity and 'Perceived recog.' represents perceived recognition as a physicist. Error bars are 95% confidence intervals on the mean.

#### Matched dataset

In what follows, we now only consider the matched dataset from the start and end of an academic session ( $n = 169$ , 110 men, 59 women). To investigate gender differences between the constructs over time we ran a series of 2 (gender: men/women)  $\times$  2 (semester: semester 1/semester 2) ANOVAs as described in the 'Analysis overview' section. We quote the estimated marginal means (M) and standard errors (S.E.) for each sample. To assess the ratio of the variances we quote the F-test. As F increases, there is increasing evidence for difference between sample means. To measure the statistical significance of this, we report the  $p$ -value. To quantify how much the samples differ we cite partial eta-squared effect sizes ( $\eta_p^2$ ). This is the proportion of variance accounted for by an independent variable, controlling for the effects of all other independent variables and interactions on the dependent variable. All effect sizes for samples with statistically significant differences were medium (.06–.013) or large ( $\geq 0.14$ ) size. Figure 2 shows the mean scores in physics identity, perceived recognition, and self-efficacy for men and women at the start and end of an academic year.

#### Physics identity

The results of our ANOVA revealed that there were significant differences between men and women with regards to their reported physics identity,  $F(1,167) = 16.50$ ,  $p < .001$ ,  $\eta_p^2 = .090$ . Examination of the estimated marginal means revealed that men ( $M = 5.39$ ,  $S.E. = .12$ ) reported greater physics identity than women ( $M = 4.56$ ,  $S.E. = .16$ ). With regards to within-person effects, there was a significant main effect of time,  $F(1,167) = 17.87$ ,  $p < .001$ ,  $\eta_p^2 = .097$ . Students reported significantly more physics identity in semester 1 ( $M = 5.19$ ,

**Table 3.** The Pearson correlations between the variables (above the diagonal) for the full dataset at timepoint 1 (309 students, 204 men, 105 women).

	ID	Self-efficacy	Perceived recognition	Grades
ID	1	.417**	.595**	.093
Self-efficacy		1	.453**	.125*
Perceived recognition			1	.186*
Grades				1

$S.E. = .11$ ) than in semester 2 ( $M = 4.76$ ,  $S.E. = .12$ ). However, there was no significant interaction between semester and gender,  $F(1, 167) = .17$ ,  $p = .683$ .

#### *Self-efficacy*

As with physics identity, there were significant differences between reported self-efficacy between men and women,  $F(1,167) = 21.83$ ,  $p < .001$ ,  $\eta_p^2 = .116$ , with men ( $M = 5.60$ ,  $S.E. = .08$ ) reporting more self-efficacy than women ( $M = 4.97$ ,  $S.E. = .11$ ). There was also a significant difference between semester 1 and semester 2,  $F(1,167) = 34.32$ ,  $p < .001$ ,  $\eta_p^2 = .170$ . The estimates revealed that students reported more self-efficacy in semester 1 ( $M = 5.49$ ,  $S.E. = .06$ ) than in semester 2 ( $M = 5.08$ ,  $S.E. = .09$ ). Again, the interaction between gender and semester was non-significant,  $F(1, 167) = .28$ ,  $p = .600$ .

#### *Perceived recognition as a physicist*

The results revealed that there was a significant gender difference with regards to students' perceived recognition as a physicist,  $F(1, 167) = 8.68$ ,  $p = .004$ ,  $\eta_p^2 = .049$ . The estimated marginal means revealed that men ( $M = 5.46$ ,  $S.E. = .10$ ) reported greater perceived recognition as a physicist than women ( $M = 4.96$ ,  $S.E. = .14$ ). There was also a significant main effect of the semester,  $F(1, 167) = 27.29$ ,  $p < .001$ ,  $\eta_p^2 = .140$ , with students reporting greater perceived recognition at semester 1 ( $M = 5.40$ ,  $S.E. = .08$ ) compared to semester 2 ( $M = 5.02$ ,  $S.E. = .10$ ). There was also a significant interaction between gender and semester,  $F(1, 167) = 5.06$ ,  $p = .026$ ,  $\eta_p^2 = .029$ . Examination of the simple main effects revealed that the difference in perceived recognition between men and women was larger in semester 2 (difference between means =  $.662$ ,  $p = .001$ ) than in semester 1 (difference between means =  $.332$ ,  $p = .046$ ). Women's perceived recognition decreased between semester 1 and semester 2, whereas men's stayed relatively constant (see figure 2).

#### *Academic performance*

The results revealed marginally significant gender differences in academic performance,  $F(1, 167) = 4.38$ ,  $p = .038$ ,  $\eta_p^2 = .026$ , with men achieving slightly higher grades than women. There was a significant main effect of the semester,  $F(1, 167) = 17.48$ ,  $p < .001$ ,  $\eta_p^2 = .095$ , with students tending to perform better in semester 2 than in semester 1. There was no significant interaction between semester and gender,  $F(1, 167) = .090$ ,  $p = .765$ .

#### *RQ2: grades as a predictor of student perceptions*

We determined Pearson correlations between the variables both in the matched sample and the full dataset from timepoint 1, in order to check global trends. Correlations for the full dataset from timepoint 1 (309 students, 204 men, 105 women) are shown in table 3. Correlations between the variables for the matched dataset (169 students, 110 men, 59 women)

**Table 4.** The Pearson correlations between the variables for the matched sample in semester 1 (above the diagonal) and semester 2 (below the diagonal).

	ID	Self-efficacy	Perceived recognition	Grades
ID	1	.471**	.675**	.239**
Self-efficacy	.516**	1	.449**	.202*
Perceived recognition	.615**	.554**	1	.326**
Grades	.196*	.294**	.287**	1

\* $p < .05$ .\*\* $p < .01$ .

are shown in table 4, with start-of-year correlations above the diagonal and end-of-year correlations below the diagonal. Tables 3 and 4 show that the correlations between self-efficacy, perceived recognition and academic performance were significant and show similar trends, with similar correlations observed in each dataset. However, the relationship between physics identity and grades was non-significant for the full sample from timepoint 1, but significant for the matched sample, which indicates that those students who answered both surveys may not be entirely representative of the full cohort.

To examine whether receiving grade information was associated with students' physics identity, perceived recognition as a physicist, and self-efficacy at a later semester (RQ2) we ran three hierarchical multiple linear regressions. In what follows  $\beta$  is the *standardized* regression coefficient, giving the predicted change in units of standard deviations for a one standard deviation change in the predictor (while controlling for the other predictors). We also quote the  $R^2$  to quantify the model fit. This is a measure of the proportion of the variance in the outcome variable that is predicted by the predictor variables.  $R^2$  will range between 0 and 1, where 0 means that the predictor variables predict no variance in the outcome variable, and 1 means the predictor variables predict all the variance in the outcome variable. This proportion is standardized such that it can also be transferred into a percentage of the variance explained, such that an  $R^2$  of .23 is the equivalent of 23% of the overall variance in the outcome variable being explained by the model.

For physics identity, we included physics identity in the first semester as a control in the regression model (Model 1, see table 5). The second regression model additionally included the grade data between semesters 1 and 2 (Model 2). The outcome variable was physics identity in the second semester (see section 'Analysis Overview' for further details of the models). Unsurprisingly, physics identity in semester 1 predicted identity in semester 2 (for full results see table 5). This was reflected in the standardized coefficient (see  $\beta_1$ , Model 1 in table 5). The addition of the grade data in Model 2 marginally improved the model fit ( $p$ -value = .048 for  $\beta_2$  in Model 2). The Delta  $R^2$  for  $\beta_2$  shows that grades predicted 1.4% of the variance in physics identity in semester 2 after controlling for physics identity in semester 1.

We ran a second hierarchical regression with the first model including perceived recognition in semester 1 predicting perceived recognition in semester 2 (for full results see table 6). The results revealed that perceived recognition in semester 1 predicted perceived recognition in semester 2, reflected in the standardized coefficient (see  $\beta_1$ , Model 1 in table 6). The inclusion of grades in Model 2 significantly improved the fit (see  $p$ -value for  $\beta_2$ , Model 2, in table 6). The Delta  $R^2$  shows that grades predicted 4.8% of the variance in perceived recognition as a physicist in semester 2, beyond that explained by perceived recognition in semester 1.

This process was repeated with self-efficacy. Self-efficacy in semester 1 predicted self-efficacy in semester 2 (for full results see table 7). The standardized coefficient was also

**Table 5.** The results of the model for physics identity and grades in semester 1 predicting physics identity in semester 2.

Outcome variable: physics identity semester 2						
Model 1: physics identity semester 1						
$Y = M + \beta_1 X_1$						
Model 2: physics identity and grades semester 1						
$Y = M + \beta_1 X_1 + \beta_2 X_2$						
Model	Predictor	Change in $F$ statistic	Degrees of freedom	Significance of change in $F$ ( $p$ -value)	Standardized regression coefficient: $\beta$	$R^2$
1 ( $\beta_1$ )	Physics identity semester 1	110.81	1,167	<.001	.632	Fit = .399
2 ( $\beta_1$ )	Physics identity semester 1	110.81	1,167	<.001	.605	Fit = .399
2 ( $\beta_2$ )	Grades	3.981	1,166	.048	.122	Delta fit = .014

**Table 6.** The results of the models, for perceived recognition as a physicist and grades in semester 1 predicting perceived recognition as a physicist in semester 2.

Outcome variable: perceived recognition semester 2						
Model 1: perceived recognition semester 1						
$Y = M + \beta_1 X_1$						
Model 2: perceived recognition semester 1 and grades						
$Y = M + \beta_1 X_1 + \beta_2 X_2$						
Model	Predictor	Change in $F$ statistic	Degrees of freedom	Significance of change in $F$ ( $p$ -value)	Standardized regression coefficient: $\beta$	$R^2$
1( $\beta_1$ )	Perceived recog. semester 1	164.21	1167	<.001	.704	Fit = .493
2 ( $\beta_1$ )	Perceived recog. semester 1	164.21	1167	<.001	.629	Fit = .493
2 ( $\beta_2$ )	Grades	17.59	1166	<.001	.232	Delta Fit = .048

**Table 7.** The results of the models for self-efficacy and grades in semester 1 predicting self-efficacy in semester 2.

Outcome variable: self-efficacy semester 2						
Model 1: self-efficacy semester 1						
$Y = M + \beta_1 X_1$						
Model 2: self-efficacy semester 1 and grades						
$Y = M + \beta_1 X_1 + \beta_2 X_2$						
Model	Predictor	Change in $F$ statistic	Degrees of freedom	Significance of change in $F$ ( $p$ -value)	Standardized regression coefficient: $\beta$	$R^2$
1 ( $\beta_1$ )	Self-efficacy semester 1	115.78	1,167	<.001	.640	Fit = .409
2 ( $\beta_1$ )	Self-efficacy semester 1	115.78	1,167	<.001	.581	Fit = .409
2 ( $\beta_2$ )	Grades	26.05	1,166	<.001	.289	Delta fit = .080

significant. The addition of grade data in Model 2 significantly improved the model fit (see table 7, the  $p$ -value for Model 2,  $\beta_2$ ). This suggests that the grades account for 8.0% of the variance in reported self-efficacy in semester 2, over and above that explained by self-efficacy in semester 1.

In summary, the results for RQ2 show that students' grades at the end of semester 1 significantly predict reported levels in semester 2 of self-efficacy, perceived recognition as a physicist, and marginal physics identity, over and above the students' reported levels in semester 1.

## Discussion

This study revealed three main findings from the matched dataset. Firstly, there are gender differences in physics identity and self-efficacy that persist across the year, with men tending to report higher scores than women (RQ1). Secondly, gender differences in perceived recognition as a physicist increased from semester 1 to semester 2 (RQ1). Finally, we found that mid-year grades predict students' perceptions of self-efficacy, perceived recognition and possibly physics identity in semester 2 (RQ2).

In more detail, the first analysis found gender differences in physics identity and self-efficacy, with men on average reporting greater physics identity and self-efficacy compared to women at both the start and end of the academic year for the matched dataset, as well as for the full dataset at timepoint 1 (the start of the academic year). Whilst the results of the Mann–Whitney U test for the full dataset at timepoint 1 did not find significant gender differences in perceived recognition as a physicist, the analysis using the matched dataset examining perceived recognition over time found significant gender differences, with the gender gap being larger at the end of the academic session than at the start. There were significant gender differences in academic performance for the full dataset at timepoint 1 and a small, yet significant main effect of gender on academic performance for the matched dataset. In the matched dataset as a whole, physics identity, self-efficacy and perceived recognition average scores were seen to decrease from the start to the end of the year.

The second analysis (to address RQ2) used several linear regressions to test whether mid-year grades predicted students' perceptions in semester 2 (the end of the academic year). The results showed that academic performance in semester 1 was associated with self-efficacy and perceived recognition in semester 2, with 8.0% of the variance in self-efficacy and 4.8% of the variance in perceived recognition explained by the mid-year grades, after controlling for self-efficacy and perceived recognition in semester 1. A more marginal result using the same analysis was found for physics identity, with 1.4% of the variance in semester 2 explained by the mid-year grades, after controlling for physics identity in semester 1. However, the analysis using the full dataset at timepoint 1 (the start of the academic year) did not find a significant correlation between physics identity and academic performance (see table 3). Thus, the regression result for physics identity is tentative, given the inconsistency in the relationship between physics identity and academic performance between the full timepoint 1 dataset and the matched dataset.

This study found that self-efficacy and perceived recognition correlated with students' grades (see tables 3 and 4) and that grades predicted students' self-efficacy and perceived recognition in the following semester (see tables 6 and 7). Kalender and colleagues [6] had suggested feedback loops between self-efficacy, perceived recognition, and grades. Our results are consistent with reciprocal correlational relationships (with the correlations suggesting associations at timepoint 1 and the regressions suggesting grades predict these factors



over time) between self-efficacy and grades, and perceived recognition and grades. However, it is important to stress that our results are only correlational and we cannot draw causal conclusions.

If such feedback loops exist, they may be particularly problematic when considering perceived recognition, given that we also found an interaction between gender and semester in the ANOVA addressing RQ1, with the perceived recognition gap between men and women widening from the beginning to the end of the academic year. We did not have the statistical power to run the regression analyses separately for men and women. This could be important future work, in terms of assessing whether grades predict perceived recognition and self-efficacy in the following semester differently for men and women. The matched dataset only spanned a single academic session and has some selection bias in terms of only including students responding to the survey at both the beginning and end of an academic session. Future work could investigate how grades correlate with each of self-efficacy, perceived recognition and physics identity over a longer time period using a fully representative sample of students.

The results from RQ1 align with previous studies on gender differences in self-efficacy e.g. [8, 9, 25] and physics identity e.g. [7], with men tending to report greater confidence in their abilities to complete physics-based tasks and greater physics identity. This study extends these findings to physics majors across all levels of the undergraduate degree in a UK context.

The tentative widening of the gender gap in perceived recognition over the academic year seen in this study may indicate the important role that instructors can play in impacting perceived recognition as a physicist [26]. Recent work has shown that a predictor of women's physics identity is whether the student feels recognized by the teacher [27, 28]. Wang and Hazari [27] highlighted that explicit and implicit attempts to make high school students feel recognized as a physicist can be internalized by students. This can take the form of explicitly telling students that they are capable of setting up tasks that make the students feel recognized without explicitly telling them. Our results indicate that this positive reinforcement and recognition of students may be particularly important across the academic session and that women, in particular, may benefit from these forms of recognition.

### Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

### ORCID iDs

Ewan Bottomley  <https://orcid.org/0000-0002-5472-2426>

### References

- [1] Marchand G C and Taasobshirazi G 2013 Stereotype threat and women's performance in physics *Int. J. Sci. Educ.* **35** 3050–61
- [2] Smith J L, Brown E R, Thoman D B and Deemer E D 2015 Losing its expected communal value: how stereotype threat undermines women's identity as research scientists *Soc. Psychol. Educ.* **18** 443–66
- [3] Platow M J, Mavor K I and Grace D M 2013 On the role of discipline-related self-concept in deep and surface approaches to learning among university students *Instr. Sci.* **41** 271–85

- [4] Smyth L, Mavor K I, Platow M J, Grace D M and Reynolds K J 2015 Discipline social identification, study norms and learning approach in university students *Educ. Psychol.* **35** 53–72
- [5] Kalender Z Y, Marshman E, Schunn C D, Nokes-Malach T J and Singh C 2019 Why female science, technology, engineering, and mathematics majors do not identify with physics: They do not think others see them that way *Phys. Rev. Phys. Educ. Res.* **15** 020148
- [6] Kalender Z Y, Marshman E, Schunn C D, Nokes-Malach T J and Singh C 2019 Gendered patterns in the construction of physics identity from motivational factors *Phys. Rev. Phys. Educ. Res.* **15** 020119
- [7] Hazari Z, Sonnert G, Sadler P M and Shanahan M C 2010 Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study *J. Res. Sci. Teach.* **47** 978–1003
- [8] Nissen J M 2019 Gender differences in self-efficacy states in high school physics *Phys. Rev. Phys. Educ. Res.* **15** 013102
- [9] Nissen J M and Shemwell J T 2016 Gender, experience, and self-efficacy in introductory physics *Phys. Rev. Phys. Educ. Res.* **12** 020105
- [10] Kalender Z Y, Marshman E, Schunn C D, Nokes-Malach T J and Singh C 2020 Damage caused by women’s lower self-efficacy on physics learning *Phys. Rev. Phys. Educ. Res.* **16** 010118
- [11] Marshman E M, Kalender Z Y, Nokes-Malach T, Schunn C and Singh C 2018 Female students with A’s have similar physics self-efficacy as male students with C’s in introductory courses: A cause for alarm? *Phys. Rev. Phys. Educ. Res.* **14** 020123
- [12] Maries A, Whitcomb K and Singh C 2021 A cross-sectional analysis of physics self-efficacy, perceived recognition, interest and identity of physics majors 2021 PERC Proc. [Virtual Conf. August 4-5, 2021] ed M B Bennett, B W Frank and R E Vieyra (<https://doi.org/10.1119/perc.2021.pr.Maries>)
- [13] Bottomley E, Mavor K I, Miles P J, Kohnle A and Wild V 2021 Gender effects in perceived recognition as a physicist and physics identity 2021 PERC Proc. [Virtual Conf. August 4-5, 2021] ed M Bennett, B Frank and R Vieyra pp 51–6
- [14] Bliuc A M, Ellis R A, Goodyear P and Hendres D M 2011 Understanding student learning in context: Relationships between university students’ social identity, approaches to learning, and academic performance *Eur. J. Psychol. Edu.* **26** 417–33
- [15] Bliuc A M, Ellis R A, Goodyear P and Hendres D M 2011 The role of social identification as university student in learning: Relationships between students’ social identity, approaches to learning, and academic achievement *Educ. Psychol.* **31** 559–74
- [16] Seyranian V, Madva A, Duong N, Abramzon N, Tibbetts Y and Harackiewicz J M 2018 The longitudinal effects of STEM identity and gender on flourishing and achievement in college physics *Int. J. STEM Educ.* **5** 1–14
- [17] Lynch D J 2010 Motivational beliefs and learning strategies as predictors of academic performance in college physics *Coll. Student J.* **44** 01463934
- [18] Sawtelle V, Brewe E, Goertzen R M and Kramer L H 2012 Identifying events that impact self-efficacy in physics learning *Phys. Rev. Special Top.-Phys. Educ. Res.* **8** 020111
- [19] Whitcomb K M and Singh C 2020 For physics majors, gender differences in introductory physics do not inform future physics performance *Eur. J. Phys.* **41** 065701
- [20] Johnson E 2013 Cronbach’s alpha ed F R Volkmar *Encyclopaedia of Autism Spectrum Disorders*. (New York: Springer)
- [21] Christmann A and Van Aelst S 2006 Robust estimation of Cronbach’s alpha *J. Multivariate Anal.* **97** 1660–74
- [22] Potvin G and Hazari Z 2013 The development and measurement of identity across the physical sciences *Proc. of the 2013 Physics Education Research Conf. (Portland, OR)* pp 281–4
- [23] Lindstrøm C and Sharma M D 2011 Self-efficacy of first year university physics students: do gender and prior formal instruction in physics matter? *Int. J. Innovation Sci. Math. Educ.* **19** 1–19
- [24] Upton G and Cook T 2014 *A Dictionary of Statistics* 2nd edn (Oxford: Oxford University Press) (<https://doi.org/10.1093/acref/9780199679188.001.0001>)
- [25] Malespina A and Singh C 2022 Gender difference in test anxiety and self-efficacy: Why instructors should emphasize low-stakes formative assessments in physics courses *Eur. J. Phys.* **43** 035701

- [26] Hazari Z and Cass C 2018 Towards meaningful physics recognition: What does this recognition actually look like? *Phys. Teach.* **56** 442–6
- [27] Wang J and Hazari Z 2018 Promoting high school students' physics identity through explicit and implicit recognition *Phys. Rev. Phys. Educ. Res.* **14** 020111
- [28] Hazari Z, Brewe E, Goertzen R M and Hodapp T 2017 The importance of high school physics teachers for female students' physics identity and persistence *Phys. Teach.* **55** 96–9