


RESEARCH

Open Access



Whose ability and growth matter? Gender, mindset and performance in physics

Alysa Malespina^{1*} , Christian D. Schunn² and Chandralekha Singh¹

Abstract

Background: Motivational factors are one active area of research that aims to increase the inclusion of women in physics. One of these factors that has only recently gained traction in physics is intelligence mindset (i.e., the belief that intelligence is either innate and unchangeable or can be developed). We studied 781 students in calculus-based Physics 1 to investigate if their mindset views were separable into more nuanced dimensions, if they varied by gender/sex and over time, and if they predicted course grade. Confirmatory factor analysis was used to divide mindset survey questions along two dimensions: myself versus others and growth versus ability aspects of mindset. Paired and unpaired *t*-tests were used to compare mindset factors over time and between genders, respectively. Multiple regression analysis was used to find which mindset factors were the best predictors of course grade.

Results: This study shows that intelligence mindset can be divided into four factors: My Ability, My Growth, Others' Ability, and Others' Growth. Further, it reveals that gender differences are more pronounced in the "My" categories than the "Others" categories. At the start of the course, there are no gender differences in any mindset component, except for My Ability. However, gender differences develop in each component from the start to the end of the course, and in the My Ability category, the gender differences increase over time. Finally, we find that My Ability is the only mindset factor that predicts course grade.

Conclusion: These results allow for a more nuanced view of intelligence mindset than has been suggested in previous interview and survey-based work. By looking at the differences in mindset factors over time, we see that learning environments affect women's and men's intelligence mindsets differently. The largest gender difference is in My Ability, the factor that best predicts course grade. This finding has implications for developing future mindset interventions and opens new opportunities to eliminate classroom inequities.

Keywords: Gender, Physics, Mindset, Motivation, Course performance

Introduction

Improving the diversity in postsecondary science, technology, engineering, and mathematics (STEM) education has been a long-standing focus of policymakers and researchers (Gonzalez & Kuenzi, 2012; Vooren et al., 2022). Physics and engineering in particular have very low numbers of women in high school courses, undergraduate programs, and in the fields (National Center

for Science & Engineering Statistics, 2019; Porter, 2019). Numerous factors affect representation in STEM fields. For example, parents of girls are less likely to believe their child could succeed in a career that requires mathematical ability (Bleeker & Jacobs, 2004; Jacobs & Eccles, 1992). Once they are in high school, girls are less likely than boys to believe that a career in physics could align with their professional goals (Hazari et al., 2010).

In physics, there are also gender disparities in introductory course performance (Kost et al., 2009; Madsen et al., 2013; Sadler & Tai, 2001; Salehi et al., 2019). Motivational factors have been linked to general academic performance (Bouffard-Bouchart et al., 1991;

*Correspondence: alm417@pitt.edu

¹ Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, USA

Full list of author information is available at the end of the article

Pintrich & De Groot, 1990; Zimmerman, 2000) as well as to gender differences in persistence in STEM courses (Goodman, 2002; Raelin et al., 2014; Sawtelle et al., 2012) and performance in STEM courses (Cavallo et al., 2004; Cwik & Singh, 2021; Marshman et al., 2017, 2018; Nissen & Shemwell, 2016).

Among the many motivational factors that have been investigated, researchers have put considerable attention on the role of intelligence mindsets (Little et al., 2019; Scherr et al., 2017). Intelligence mindset describes a person's beliefs about the nature of intelligence: is it innate and unchangeable or something that can be developed with effort (Dweck, 2006)? In more recent years, a focus has shifted to discipline-specific intelligence mindsets since students appeared to have separate views by discipline and the disciplines-specific mindset was more predictive of student performance in the discipline (Kalender et al., 2022; Marshman et al., 2017). However, since physics-specific mindset is still a very recently explored concept, many fundamental questions about its nature and relationship to gendered performance in physics are still open. Specifically, we address the research questions:

- RQ1. What are the components to students' physics intelligence mindsets?
- RQ2. Are there gender/sex differences in the components of students' physics intelligence mindsets?
- RQ3. If there are differences in the components of students' physics intelligence mindsets, do the differences grow or decline from the beginning to the end of their first university-level physics course?
- RQ4. Do any of the mindset components predict course grade?

To answer these questions, we chose the first calculus-based introductory course as the research context. Introductory calculus-based physics courses are typically taken by engineering and physical science majors, while most algebra-based physics students are life science and pre-medical majors. As a result, calculus-based introductory physics courses are likely to be majority men, which likely further reinforces stereotypes and negative messages that women in physics courses are receiving (Miller et al., 2015; National Center for Science & Engineering Statistics, 2019; Porter, 2019). Because of the inequities in these courses and the underrepresentation of women, finding effective ways to measure and improve physics mindset is particularly important in this population if we wish to make physics classrooms more equitable for women and gender minorities. If physics mindset is a useful predictor of learning outcomes, then improving it in physics students may help overall outcomes and equity

of outcomes in engineering, physics, and other physical science fields.

Theoretical background

Intelligence mindset theory

Carol Dweck and her colleagues theorized two types of intelligence mindset—growth and fixed—in the late twentieth century. A growth mindset is one in which intelligence is viewed as something that can be cultivated with effort, like a muscle, whereas a fixed mindset is one in which intelligence is thought to be innate and unchangeable (Dweck, 2006). In the original conception, researchers conceived intelligence mindset as a single continuum in which students varied from having a strong growth mindset at one end of the continuum to having a strong fixed mindset at the other end of the continuum. However, in recent years researchers have used both continuum models (see Yeager & Dweck, 2020) and models with separable dimensions in which students can endorse both (or neither) simultaneously (Cook et al., 2017; Shih, 2011; Troche & Kunz, 2020). The original view holds that as a student ceases to endorse a fixed mindset, they will necessarily endorse a growth mindset (Yeager & Dweck, 2020). In a two-factor model, it may be possible for a student to endorse neither growth nor fixed beliefs, or they may endorse both types of beliefs. For example, a student might think some basic foundational intelligence or talent is required in addition to seeing value in practice towards further developing intelligence.

The mindsets held by a learner are thought to shape how students engage in learning. With a fixed mindset, a student will disengage from or avoid difficult tasks; with a growth mindset, a student will view struggle as an opportunity to learn and gain skills, and therefore will welcome such challenges (Muenks & Miele, 2017; Yeager & Dweck, 2012). Six-year-old girls are more likely to say that boys are “really, really smart” than they are to say girls are, and to avoid activities that are said to be for children who are “really, really smart” (Bian et al., 2017). The engagement, propensity to attempt challenging problems, and persistence that come with a growth mindset have been linked to positive learning outcomes (Dweck, 2007, 2008; Limeri et al., 2020), even after controlling for prior academic achievement (Aronson et al., 2002; Blackwell et al., 2007; Good et al., 2003).

Intelligence mindsets may also play a role in shaping learner self-efficacy (Bandura, 1997, 2012) and in experiences of anxiety in learning and testing environments (Bandura, 1997; Zeidner, 1998). As a result, growth mindsets are not only relevant to improving learning outcomes for all students, but they also may be an important factor in creating equitable classroom environments. For example, having a growth mindset has been linked to

greater participation in STEM fields, especially for students from racial and ethnic underrepresented groups (Kricorian et al., 2020; Rattan et al., 2015). Additionally, both students in underrepresented groups and women reported a greater sense of belonging if they endorsed a growth mindset (Rattan et al., 2018). Growth mindsets can be particularly useful for students as a way to combat stereotype threat. Stereotype threat is “being at risk of confirming, as self-characteristic, a negative stereotype about one’s group” (Steele & Aronson, 1995, p. 797). For example, a girl or woman taking a math test may feel anxious because of cultural stereotypes that women are not as good at math as men. When such stereotype threats are combined with a fixed mindset, withdrawal from efforts in mathematics can result: the student cannot change their gender, race, or culture, so they may choose to divest from a field that leaves them anxious about representing these identities poorly (Steele & Aronson, 1995).

Although intelligence mindsets are carried by students into various learning contexts (i.e., have some stability over time and context), they can be malleable through strategic (and relatively brief) interventions with positive results for students’ learning outcomes, such as mathematics assessments (Bagès et al., 2016), standardized test outcomes (Good et al., 2003), and course grades (Blackwell et al., 2007) especially if students are at a high risk of failing a class (Yeager et al., 2016, 2019). However, the effectiveness of both mindset as a predictor of student success as well as the methodology and effectiveness of mindset interventions has been found to vary greatly (Denworth, 2019; Sisk et al., 2018). For example, only 12% of the interventions included in a recent meta-analysis resulted in significantly greater academic achievement (Sisk et al., 2018), which may make some instructors concerned about their use of class time (Hattie, 2012).

The Sisk et al. study explores several potential reasons the effectiveness of these interventions varies, they tend to focus on technical (i.e., if the intervention is online or in-person, the length of the intervention, etc.) differences, which may not be the only aspects of importance. Yeager and Dweck (2020) offer more explanations of the varied effectiveness of mindset interventions: first, they show concern about moderation of an intervention’s effectiveness at the study level (for example, by length of intervention) rather than the student level (for example, by student gender or socioeconomic status), as it can be difficult to discern the effectiveness of an intervention without simultaneously knowing of methods of the intervention, the students who receive the it, and the larger context the intervention takes place in (e.g., if a growth mindset is supported in the classroom after the intervention). There is also concern about the procedural

differences among mindset intervention studies: for example, an intervention that simply explains what a growth mindset is will not be as effective as one that offers students concrete actions to utilize such a mindset (Yeager & Dweck, 2020).

We also hypothesize that some of the varying effectiveness of the interventions may be due to procedural details in the interventions. One possibility is that intervention effectiveness relies on customization to the particular concerns that students have in a particular context. Another possibility is that the focus of the intervention affected its outcome. For example, did the intervention seek only to address the growth mindset but ignore the ability mindset?

Another conceptual divide in mindset research involves beliefs about self versus others. De Castella and Byrne (2015) found that Australian high-school students conceptualized intelligence mindsets differently for themselves than for others. They also found that intelligence “self-theory” was a stronger predictor of academic performance and motivation than general intelligence mindsets. Some prior interventions have tried to convince students that people in general can grow their intelligence, leaving relatively untouched the beliefs they have about themselves.

A third issue might also exist in domain-specificity of intelligence mindsets. That is, students might believe that intelligence in general can change through hard work but still have fixed mindsets about particular domains that then more strongly shape how they engage in those particular domains. For example, it was physics-specific mindsets rather than general intelligence mindsets that predicted performance in physics classes (Marshman et al., 2017). Further, many stereotypes about women and students from underrepresented racial and ethnic groups (for example, Black or Latinx students) are highly domain-specific (e.g., strengths in arts and humanities, weaknesses in math and sciences; Eaton et al., 2020; Ganley et al., 2018). Indeed, women in general have higher grades on average in high school and in university (Voyer & Voyer, 2014), so a domain-specific mindset would make more sense as contributing to performance differences in physics courses.

Physics intelligence mindsets

There appear to be common views both in society and within the discipline that physics requires a special brilliance. In a study of brilliance beliefs by academic discipline, physics faculty, post-doctoral researchers, and graduate students were more likely to say that physics requires innate talent than those in almost all other fields (Leslie et al., 2015). Brilliance beliefs are not the same as a fixed mindset, though they work in tandem. If a student

thinks raw talent is needed to succeed in a domain (a brilliance belief), and they believe that intelligence is unchangeable (a fixed mindset), then they will see no path to success unless they believe they have innate talent (Deiglmayr et al., 2019). Indeed, the Leslie et al. (2015) brilliance study revealed a negative correlation between degree of endorsement in ability beliefs and percentage of PhDs who are women or are from underrepresented racial and ethnic groups, with physics being second highest (after mathematics) among STEM disciplines in field-specific ability beliefs and lowest in percentage of women with doctorates. In a recent study, only half of graduate admissions committees in physics prioritized a growth mindset in their selection process, meaning that they prioritized potential for growth, rather than exclusively seeking out the students with the highest grades and Graduate Record Examinations (GRE) scores (Scherr et al., 2017).

Physics-specific mindset research has just begun in recent years. Interviews show that students (Little et al., 2019) and faculty (Scherr et al., 2017) may simultaneously endorse both growth and fixed mindset beliefs, pointing to a need for nuanced measures of mindset. Meanwhile, survey data have provided evidence that students' physics mindsets can be different than their general intelligence mindsets (Marshman et al., 2017). But an open question regarding physics-specific mindset involves its nature. In particular, does it also separate into independent dimensions of growth and fixed mindsets, with students independently endorsing or denying fixed (fundamental talent) and growth (ability to further improve) aspects? We turn to this issue and potential dimensions of physics mindset in the next section.

Dimensions of physics intelligence mindsets

As noted previously, prior research on general intelligence mindsets has gradually transitioned from strict characterization as one continuum (with fixed and growth mindsets on either end Blackwell et al., 2007; Dweck, 2006) to considering a two-factor model measuring endorsement of fixed and growth mindsets separately (Cook et al., 2017; Shih, 2011; Troche & Kunz, 2020), which we denote as growth versus ability because the fixed label seems to connote the absence of growth rather than the presence of a foundational talent. The primary evidence in favor of treating them separately as two dimensions is psychometric evidence in which a two-factor model produced a better fit to the data. To date, evidence supports a separation of growth and ability dimensions in physics mindsets as well (Kalender et al., 2022), although some researchers have applied a single dimension approach to their data (e.g., Kepple et al., 2020; Marshman et al., 2017).

Another divide which has recently emerged in mindset research is the me versus others distinction. As noted earlier, De Castella and Byrne (2015) found that intelligence about the self was the stronger predictor of academic performance. A recent study about physics intelligence mindsets, Kalender et al. (2022) found that physics intelligence mindsets divided into four components along the combinations of me versus others and growth versus ability. Although the four components showed some correlations with each other, the best fitting model to the survey data separately measures the four components: My Ability (students' beliefs about their own abilities), My Growth (students' beliefs about their own potential to grow), Others' Ability (students' beliefs about others' abilities), and Others' Growth (students' beliefs about others' potential to grow). Further, the My Ability component was the best predictor of physics course grade, had the largest gender differences, and appeared to largely mediate the effects of gender on grades.

However, the Kalender study uncovered the four physics intelligence mindset components through exploratory quantitative analyses of survey results from a survey that was not designed to separately measure four components of physics intelligence mindsets. There were only one or two survey items for each component's measure, and the items also sometimes differed in other ways across components. In this study, we aim to expand on these results using a larger set of survey items that were specifically designed to measure these four components, allowing for more robust test of the separation into these four components, as well as replicate the three main findings from that study: My Ability was found to be the main predictor of course grades, it was also the component showing the largest gender difference, and it was found to be the only component that mediates the relationships between gender and grades.

Materials and methods

Participants

This study takes place at University of Pittsburgh, a large (19,017 degree-seeking undergraduates in 2019), public, urban, predominantly White institution (78% in 2019; University of Pittsburgh, 2021) in the northeastern United States. The university is very high research activity, doctoral, and has an acceptance rate of 57% (Integrated Postsecondary Education Data System, 2020). The participants were students enrolled in calculus-based Physics 1 over one semester and across four course sections, each taught by a different instructor. The course covers mechanics and waves, and is taught in a traditional lecture-based format. The $N=683$ students included in the study were those who completed at least pre- or post-surveys and passed an attention check (a question

inserted in the survey that requested students answer “C”). Some ($N=39$) students were excluded from some portions of this study because they were missing either course grades or prior academic preparation information, though these students were included in the survey validation portion of this study.

Demographic data were acquired from university records. In the student sample, 63% were enrolled in the college of engineering, and virtually all of these engineering students (99%) were in their first semester at the university. The rest of the students were primarily science majors and 59% were from later years at the university. Based upon the data available from the university, women constituted 36% of the student sample. According to university-provided race/ethnicity data, students identified as follows: 73% White, 13% Asian, 7% Hispanic/Latinx, 4% multiracial, 2% African American/Black, and 1% unspecified.

Measures

Demographic information

Students provide demographic information as part of university enrollment. Students were given the binary options “male” and “female” to identify their gender upon entering the university, although this conflates gender and sex (Schudson, 2021). We acknowledge the harm that such data collection practices cause (Traxler et al., 2016), and we are pleased to report that our university has recently switched to collecting gender information using more than binary options. Given the limitations of the data source, the patterns will predominantly reflect patterns of cis-gendered women and men. This approach marginalizes non-binary and other gender minority students (Traxler et al., 2016; Van Dusen & Nissen, 2020). However, we use the data collected by the university (i.e., the options provided were female and male while labeled as gender) and refer this variable as “Gender/Sex” in our analysis and results sections (Schudson, 2021).

For the quantitative analyses, gender/sex was coded as an indicator variable: women = 1, men = 0. For race and ethnicity, students were given six options (American Indian or Alaska Native, Asian, Black or African American, Hispanic or Latino, Native Hawaiian/Other Pacific Islander, and White), and students could choose multiple options. Race/ethnicity was only used in description of the sample, and was coded as a series of indicator variables for each major category (i.e., included that racial/ethnic identity or not).

Physics intelligence mindset

We adapted this mindset survey from previously validated surveys (Kalender et al., 2022). The survey was designed to measure mindsets about self and others, as

well as growth- and ability mindsets. Therefore, to be able to separately assess these different aspects of mindset, additional questions were created and some questions were adapted to make the more specific focus more salient. For example, “People can change their intelligence in physics quite a lot by working hard”, becomes “I can change my intelligence in physics quite a lot by working hard.” After the questions were drafted, we used semi-structured cognitive interviews to ensure that students interpreted questions as intended. We conducted 20 one-hour interviews with students who had previously taken physics courses ranging from introductory to graduate-level. Participants were compensated \$25. We oversampled on women, given the research focus. A few questions were edited slightly after the interviews, and the questions in the final set were generally interpreted as intended. The final survey had 19 items, each on a four-point Likert scale (*Strongly Disagree, Disagree, Agree, Strongly Agree*): seven My Ability items, four My Growth items, four Others’ Ability items, and four Others’ Growth items. See Appendix A for the full set of items. For analysis, the four ratings levels were recoded as 1 to 4, with reverse coding for all my ability and others’ ability questions (e.g., questions 5–11 and 16–19). Prior Rasch modeling (Frey, 2018) with this four-point scale for mindset items had found roughly equal psychological distance between levels, justifying use of mean scores (Kalender et al., 2022; Marshman et al., 2017).

Prior academic preparation

Two measures of prior academic preparation were used as control variables in the analyses. High school Grade Point Average (HS GPA) was reported using the weighted 0–5 scale, which is based on the standard 0 (Failing)–4 (A) scale with adjustments for Honors, Advanced Placement and International Baccalaureate courses (all of these programs may offer a “weighted” GPA that adds up to one grade point as a reward to taking advanced courses, which can allow a GPA higher than 4.0). Approximately 1% of students had high school GPAs over 5. They were excluded from the study because their high schools likely used a different grading system. HS GPA is regularly found to predict early undergraduate course performance and is taken to be a measure of general academic skills related to self-regulation, attendance, and putting effort into assignments (Galla et al., 2019). Students’ Scholastic Achievement Test math (SAT math) scores were used as proxies for mathematical problem-solving skills at the time of university admission. SAT math is one predictor of college performance (Galla et al., 2019), particularly in quantitative courses like introductory physics (Crisp et al., 2009; Hazari et al., 2007; Vincent-Ruz et al., 2018). The scores are on a scale of 200–800. We mediated

outliers in SAT math by winsorizing (Frey, 2018). To winsorize the scores, we replaced outliers with values two standard deviations above or below the mean, so that we maintained the direction of the outlier without introducing extreme values. If a student took the American College Testing (ACT) examination, we converted ACT to SAT scores (The College Board & ACT, Inc., 2018). SAT scores had a negative skew. If a student took a test more than once the school provided the highest section-level score if a student took the SAT and the highest composite score if the student took the ACT. If a student took both tests, we used their SAT score.

Physics course grade

Physics 1 course grades, the primary course performance measure, were based on the 0–4 scale used at this university, with A=4, B=3, C=2, D=1, F=0 or W (late withdrawal), where the suffixes ‘+’ and ‘-’, respectively, add or subtract 0.25 grade points (e.g., B- = 2.75 and B+ = 3.25), except for the A+, which is also reported as 4. Each course instructor determined their own grading schemes and there was not a shared departmental exam. However, from examination of syllabi across all sections, course grades were predominantly based upon traditional midterm and final exams, with a smaller portion of the grades based on homework, quizzes, and recitation attendance. Course grades had a negative skew.

Procedures

The surveys were administered to students during recitations associated with the course. The 50 min recitation sections are mandatory and led by teaching assistants (TAs). The first (“pre”) survey was administered on paper during the first or second week of classes, and the final (“post”) survey was administered last week of classes. The mindset items were a subset of a larger survey, which took approximately ten minutes to complete. To encourage a high completion rate, students receive either a participation grade or a small amount of extra credit for completing the survey, depending on the instructor’s preference. 80% of course enrollees completed the survey at pre and 41% did so at post, reflecting a lower recitation participation at the end of the semester. However, the student sample that completed the survey is very similar to the general population of the course in terms gender/sex, prior preparation, and course performance, and the students that completed only the pre-survey are similar to those who took both the pre- and post-surveys (see Additional file 2: Appendix B).

Survey results were collected, de-identified by an honest broker, and then combined with similarly de-identified demographic information and academic history.

Analyses

Survey validation

Confirmatory factor analysis (CFA) using the R package “lavaan” was used to both provide quantitative validation of the survey items and to test the proposed conceptual division into four components in terms of growth/ability and myself/others. To evaluate if the model was acceptable, we chose the following standards: standardized factor loadings of each item were all above 0.5 (Kline, 2016, p. 301), a Comparative Fit Index (CFI) and Tucker–Lewis index (TLI) greater than or equal to 0.95 (Hu & Bentler, 1999), a Root Mean Square Error of Approximation (RMSEA) less than or equal to 0.05 for “good fit” or 0.08 for “fair” fit (Browne & Cudeck, 1992), and a Standardized Root Mean Square Residual (SRMR) less than or equal to 0.06 (Hu & Bentler, 1999). The survey was designed to divide items into four categories, but we also explored if a one-factor or two-factor model (i.e., only dividing along one of the aforementioned dimensions, rather than ability/growth and myself/others simultaneously) resulted in a better fit. Poorly fitting items were dropped, and the model was re-evaluated with the remaining items. To create latent variables, we calculated the average score of the questions in each validated category. All the mindset factors are scored from 1 to 4, and are coded such that a high score corresponds with a growth/malleable physics mindset, and a low score corresponds with a fixed mindset. After averaging scores, we winsorized each mindset factor so that outliers were set at a cutoff two standard deviations from the mean of each factor. The resulting variables were used as the mindset measures for the rest of the study.

Descriptive statistics

To characterize change in mean attitudes over time, and differences by gender/sex in mean attitudes at pre and post as well as grades, we used Cohen’s *d* to describe the size of the means differences and *t*-tests to evaluate the statistical robustness of the differences. Cohen’s *d* is considered small if $d \sim 0.2$, medium if $d \sim 0.5$, and large if $d \sim 0.8$ (Cohen, 1988). Paired *t*-tests were used to compare mindset factors between pre and post, while unpaired *t*-tests were used to compare mindset factors between genders/sexes. Levene’s test (Frey, 2018) was implemented to ensure that the homogeneity of variance assumption was met for the unpaired *t*-tests. We used a significance level of 0.05 in the *t*-tests and the later regression models as a balance between Type I (falsely rejecting the null hypothesis) and Type II (falsely accepting the null hypothesis) errors (Frey, 2018). The change-over-time analyses were also done

for all instructors separately to check for consistency of the patterns across instructors.

Pearson correlations were calculated between the generated latent variables and between the pre- and post-survey scores of the same variable. These correlations provided information on potential problems of collinearity among predictors in the multiple regressions (e.g., Pearson $r > 0.7$). Further, Pearson correlations allowed us to examine attitude stability over time during this first experience with university-level physics. We also found correlations between mindset factors and course grades as a baseline prediction model.

Predicting learning outcomes

Multiple linear regression analysis was used to find partial correlations between mindset factors and grades, controlling for gender/sex and prior preparation. We chose to use regression analysis instead of hierarchical linear modeling because we find the Interclass Correlation Coefficients of motivational factors in these courses are regularly smaller than 0.04 and always smaller than 0.10. Multiple models were tested in order to find which was the best predictor of learning outcomes and show robustness of relationships across model specification. All models used standardized regression coefficients as a measure of effect size. The models were implemented using Stata statistical software (StataCorp, 2021). To test the normality of errors, we compared a kernel density estimate of each model's residuals with a normal distribution. Each model had a normal distribution of residuals.

A baseline model predicted grade using only gender/sex, high school GPA, and SAT math scores. Next, we added the mindset variables with the strongest correlation to grade one-by-one until all mindset variables were present. All models with significant mindset variables were kept, along with the final model with all variables induced as a robustness test.

The regression analyses were repeated with two sets of attitudinal variables: first the scores from the pre-survey, then from the average of pre- and post-survey scores. The average group included only students who took the survey both times. Average scores were used instead of post-survey scores for two reasons. First, post-survey scores raise the question of causality (did course performance affect mindset or did mindset affect course performance?). Second, the average score is a proxy for students' mindset during the semester, while they were taking the course, rather than after the class. Using average rather than only pre-survey data is particularly important given the sizeable changes from pre to post that were observed in several of the attitudinal variables.

Results

RQ1: What are the components to students' physics intelligence mindsets?

One of the 19 survey items ("I will always be as good at physics as I was in high school.") was removed as a first step because the cognitive interviews show that students did not interpret it as intended. All other survey items appeared to be interpreted as intended.

Five additional survey items were removed during the CFA model testing process due to consistently low factor loadings or cross-loading that led to a poor overall model fit. The removed items are indicated in italics on the full survey shown in Additional file 1: Appendix A. Of the four tested models (using all questions in a single factor; splitting by a "growth/ability" dimension; and splitting by a "myself/others" dimension; dividing into four categories in the combination of both dimensions), both two-category models were rejected, as they failed to meet accepted cutoff values for our chosen fit indices. The third model, which divided survey items into four categories and can be seen in Table 1, had a good model fit. The CFI was 0.95, the TLI was 0.95, the RMSEA was 0.073, and the SRMR was 0.052. Three of the fit indices—CFI, TLI, and SRMR meet our chosen cutoffs. Our RMSEA meets Browne and Cudeck's (1992) ≤ 0.08 guideline for acceptable fit. All standardized factor loadings were above a 0.50 threshold. We named the resulting categories "My Ability" (MA), "My Growth" (MG), "Others' Ability" (OA), and "Others' Growth" (OG). Three categories—MA, MG, and OG—had acceptable values of internal consistency (Cronbach $\alpha > 0.7$), while OA had slightly lower reliability ($\alpha = 0.68$). All four categories had negative skew (e.g., a skew toward a growth mindset). To confirm that these factors held equally well for men and women, we performed measurement invariance testing and found that both weak and strong invariance held for these factors (see Additional file 3: Appendix C). Table 2 displays Pearson correlations between the four mindset groups. Intercorrelations among the scales are all moderate and positive (after reverse coding of ability), but none are so high as to represent redundant measures. There is also not strong organization of these correlations at the level of the dimensions: while there are some pairwise combinations that are higher, on the whole there are four scales that are all moderately correlated with one another.

All four factors show moderate stability over time. Thus, the attitudes that students had at the beginning of the semester could have provided the opportunity to continuously influence student performance and behaviors during the whole semester. However, because there is also significant change, the average attitude held across the semester is likely a better estimate of the relationship of attitudes to performance.

Table 1 Survey items included in the study

Survey item text	λ
My growth ($\alpha = 0.84$)	
I can become even better at solving physics problems through hard work	0.76
I am capable of really understanding physics if I work hard	0.83
I can change my intelligence in physics quite a lot by working hard	0.82
My ability ($\alpha = 0.84$)	
Even if I were to spend a lot of time working on difficult physics problems, I cannot develop my intelligence in physics further	0.64
I won't get better at physics if I try harder	0.64
I could never excel in physics because I do not have what it takes to be a physics person	0.87
I could never become really good at physics even if I were to work hard because I don't have natural ability	0.87
Others' growth ($\alpha = 0.84$)	
People can change their intelligence in physics quite a lot by working hard	0.82
If people were to spend a lot of time working on difficult physics problems, they could develop their intelligence in physics quite a bit	0.82
People can become good at solving physics problems through hard work	0.77
Others' ability ($\alpha = 0.68$)	
Only a few specially qualified people are capable of really understanding physics	0.67
To really excel in physics, people need to have a natural ability in physics	0.73
If a student were to often make mistakes on physics assignments and exams, I would think that maybe they are just not smart enough to excel in physics	0.55

The Survey Item Text column contains factor names and internal consistency using Cronbach's α , in addition to the survey item text. The λ column contains standardized factor loadings, using both pre- and post-survey results ($N = 781$)

Table 2 Pearson correlations among attitudinal measures

	Pre				Post			
	MG	MA	OG	OA	MG	MA	OG	OA
Pre-survey intercorrelations								
Pre-MG								
Pre-MA	0.52							
Pre-OG	0.51	0.43						
Pre-OA	0.28	0.45	0.33					
Pre-post intercorrelations					Post-survey intercorrelations			
Post-MG	0.34							
Post-MA		0.44			0.67			
Post-OG			0.38		0.69	0.57		
Post-OA				0.33	0.43	0.65	0.44	

Clockwise from top left: pre, post, and pre- versus post-Pearson correlations between constructs. The following terms are abbreviated: My Growth (MG), My Ability (MA), Others' Growth (OG), and Others' Ability (OA). All correlations are statistically significant at $p < 0.001$

RQ2: Are there gender/sex differences in the different components of students' physics intelligence mindsets?

Table 3 shows descriptive statistics for each measure by gender/sex at pre and post. On the pre-surveys, men and women have nearly identical and high scores in the My Growth, Others' Ability, and Others' Growth categories. That is, in general most students have growth rather than fixed mindsets, particularly when considering others. The only pre-survey category with a significant gender/sex difference is My Ability. In this category the gender/sex difference has a medium effect size with men having

higher scores than women (i.e., women were more likely than men to believe that natural ability is important for themselves to succeed in physics).

There were also gender/sex differences in prior academic performance. As seen in Table 3, women tend to have higher high school GPAs than men in our sample, but lower SAT math scores. Both of these differences had relatively small effect sizes and both populations generally had high scores (i.e., were generally well prepared for challenging academic work). Thus, the lower average course grades for women (see Table 3) are somewhat

Table 3 Descriptive statistics and gender/sex differences for all measures

Variable	Men			Women			Gender/sex contrast	
	N	Mean	SD	N	Mean	SD	t	d
My growth								
Pre	307	3.62	0.48	190	3.59	0.46	0.57	0.05
Post	161	3.48	0.54	84	3.12	0.58	4.76***	0.64
My ability								
Pre	307	3.49	0.46	190	3.32	0.46	3.99***	0.37
Post	161	3.28	0.62	84	2.98	0.59	3.69***	0.50
Others' growth								
Pre	307	3.50	0.49	190	3.48	0.49	0.47	0.04
Post	161	3.45	0.54	84	3.17	0.44	3.93***	0.53
Others' ability								
Pre	307	3.16	0.55	190	3.15	0.53	0.13	0.01
Post	161	3.08	0.61	84	2.91	0.49	2.24*	0.30
SAT Math	342	712	60	203	695	64	3.12**	0.28
HS GPA	342	4.09	0.42	203	4.25	0.35	-4.45***	-0.39
Course Grade	342	2.76	0.98	203	2.49	0.90	3.07**	0.27

Mean and standard deviation (SD) by gender/sex of each mindset factor at pre and post, along with SAT Math, HS GPA, and Physics 1 grade, and Cohen's *d* and *t*-test of gender/sex differences. Positive values of *d* and *t* indicate that men have a higher score

p* < 0.05, *p* < 0.01, ****p* < 0.001

surprising from an academic preparation perspective. Note, we cannot assume men's higher SAT math scores directly translate into higher grades in math-intensive courses. There is no similar gendered grade difference in Calculus 1, which this population often takes in tandem with Physics 1 (Whitcomb & Singh, 2020; Whitcomb et al., 2021). Instead, factors other than academic preparation are likely at play.

RQ3: If there are differences in the components of students' physics intelligence mindsets, do the differences grow or decline from the beginning to the end of their first university-level physics course?

By the end of the semester, there were moderate-to-large gender/sex differences in all four mindset constructs, and all gender/sex contrasts became statistically significant. Thus, following their first experience in university-level physics, women were more likely than men to believe that natural ability is important to succeed in physics for both themselves and others. This change in gender/sex differences reflects moderate-to-large declines in attitudes in women but only small declines in men, on average (see Fig. 1). This suggests that classroom experiences that influenced student mindsets affected men and women differently. Trends were similar across instructors, though some results were nonsignificant when calculated for individual instructors' classes, due to low sample size.

RQ4: Do any of the mindset factors from RQ1 predict course grade?

We conducted multiple regression analysis to find which of the four mindset factors best predicted physics course grade (see Table 4). Models 1–3 used only pre-survey results, while Models 4–6 used the mean of pre- and post-survey mindset scores (because of the large changes in mindset across the semester). In Model 1, only gender, SAT math scores, and HS GPA are included as predictors and all three were statistically significant. This model shows that women have lower Physics 1 grades than men when controlling for prior academic preparation, formally establishing that other factors are needed to account for gender/sex differences in course performance.

Model 2 includes My Ability (MA) as a fourth predictor, the single strongest correlate of grades. Here pre-survey MA is a significant predictor beyond academic preparation. Its addition weakens the relationship between gender/sex and Physics 1 grade, though gender/sex remains significant. Additionally, Model 2 has a small increase in adjusted R-squared compared to Model 1. This means that Model 2 explains more of the variance in course grades than Model 1, while penalizing for nonsignificant predictors (Frey, 2018). Model 3 adds the remaining pre-survey mindset factors: MG, OA, and OG. None of the newly added factors are statistically significant, and their addition leaves fully intact or slightly strengthens the predictive power of the other predictors,

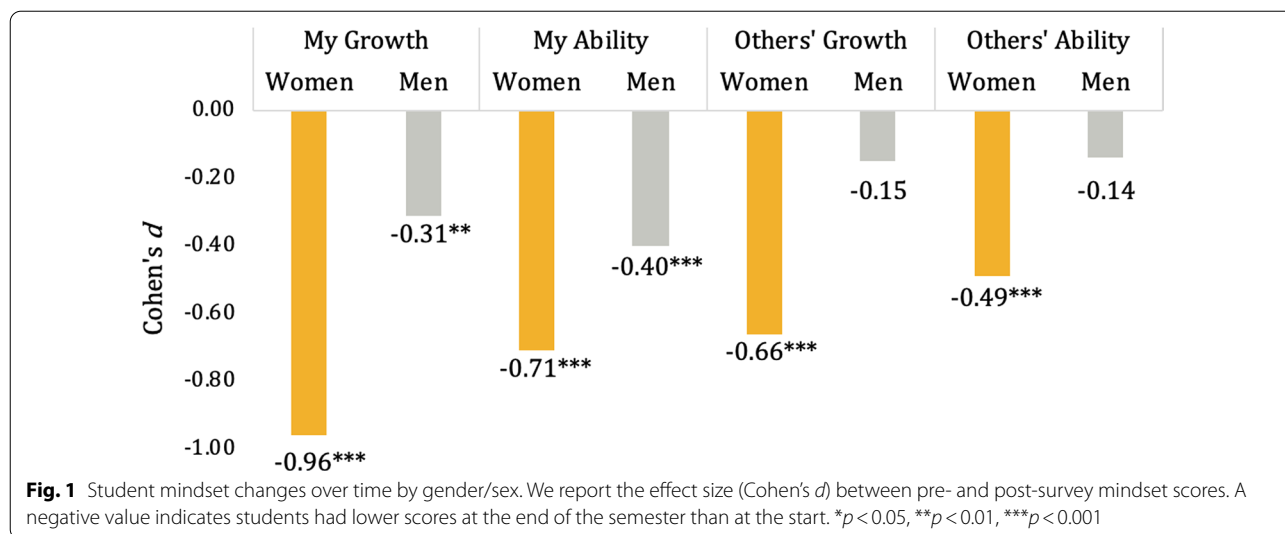


Table 4 Regression models predicting physics course grades

Variable	Pre			Pre/post average		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Gender/sex (W = 1, M = 0)	-0.13***	-0.12***	-0.11**	-0.15*	-0.09	-0.09
HS GPA	0.32***	0.32***	0.32***	0.28***	0.28***	0.28***
SAT math	0.39***	0.38***	0.38***	0.34***	0.31***	0.31***
My growth			0.03			0.07
My ability		0.09*	0.10*		0.25***	0.24***
Others' growth			-0.01			-0.02
Others' ability			-0.05			-0.06
N	497	497	497	197	197	197
Adjusted R ²	0.32	0.33	0.33	0.27	0.33	0.32

This table shows standardized regression coefficient (β) values of multiple regression analysis predicting Physics 1 course grade using pre-survey responses (Models 1–3) and an average of pre- and post-survey responses (Models 4–6). Adjusted R² and N are reported for each model

p* < 0.05, *p* < 0.01, ****p* < 0.001

suggesting robust relationship estimates. The predictive power of gender/sex decreases slightly. Variance Inflation Factors (VIFs) for every variable in Models 1–3 were below our cutoff of 2.0, which indicates that our models are not skewed by multicollinearity, even in the case of the different mindset factors that were also moderately correlated with one another.

Models 4–6 are focused on the sample that completed both pre and post to unpack the predictive role of average attitudes across pre and post. Model 4 is identical to Model 1, but now providing the baseline model for the reduced sample set. The parameter values are similar in approximate magnitude as those of Model 1, although the SAT estimate is smaller and the gender/sex estimate is larger.

Model 5 adds average MA as a predictor. Average MA has more than twice the predictive power of pre-MA, and

the gender/sex estimate decreases in size by 40%. Model 6 introduces the remaining average mindset factors, none of which are statistically significant predictors, similar to the findings of Model 3. There are no major changes in the predictive power of MA, HS GPA, or SAT math from Model 5 to Model 6, again suggesting robust relationship estimates and that MA in particular was the most likely mediator of gender/sex differences in grades among the mindset factors.

In Models 4–6, VIFs are mostly below the cutoff of 2.0, except for MG (VIF=3.08) and MA (VIF=3.19) in Model 6. MA and MG are often conceptualized as a single factor (García-Cepero & McCoach, 2009) because they have substantial intercorrelations (Troche & Kunz, 2020) as in our analysis (see Table 3). However, the robustness of the pattern of regression estimates and much lower predictive power of MG across models

supports the focus on MA as the key predictor of student performance.

Although there was analytic support for treating the Likert ratings as continuous predictors, some skew in the distributions did occur. However, regression results were very similar when binary mindset variables (i.e., 1 for strong endorsement of growth mindset/strong rejection of fixed mindset; 0 otherwise) were used instead of means based upon 1–4 codings. Most importantly, MA was the strongest predictor of grade among the mindset factors.

Discussion

RQ1: What are the components to students' physics intelligence mindsets?

The current study strongly replicated the exploratory findings of Kalender et al. (2022) using a survey instrument designed to specifically test for the four components: My Ability (MA), My Growth (MG), Others' Ability (OA), and Others' Growth (OG). It also builds upon the work of De Castella and Byrne (2015), who found an empirical separation of my versus others' mindset factors, along with a number of other studies that found support for a divide along the ability/effort dimensions (Cook et al., 2017; Shih, 2011; Troche & Kunz, 2020). The four components were only moderately correlated with one another (~25% shared variance at pre) and were separable in CFA models. Further, the My Ability factor showed different patterns related to RQ2 and 3. In sum, there was support for separating our four components both from psychometric analyses and empirical phenomena.

RQ2: Are there gender/sex differences in the different components of students' physics intelligence mindsets?

At the start of the semester, there were no gender/sex differences in My Growth, Others' Growth, or Others' Ability. However, there was an initial (moderately sized) gender/sex difference in My Ability even among this relatively selective set of students who have opted into engineering and physical science pathways. That is, women in this context were more likely than men to believe that physics requires innate ability and that they, in particular, did not possess that ability. However, by the end of the semester, all four mindset categories showed significant gender/sex differences, and sometimes large differences.

RQ3: If there are differences in the components of students' physics intelligence mindsets, do the differences grow or decline from the beginning to the end of their first university-level physics course?

Both self-theory mindset factors (My Ability and My Growth) significantly decreased (i.e., mindsets became

less growth-oriented and more fixed) for men from the start to the end of the semester, while all intelligence mindset factors significantly decreased for women. In addition to decreasing all mindset factors for students regardless of gender, the courses also created or contributed to a gender-based inequity in physics intelligence mindsets.

These results add to research showing that women in physics courses also have other forms of lower average motivational characteristics, such as self-efficacy and sense of belonging, than do men, even in highly self-selected pathways (Marshman et al., 2017; Nissen & Shemwell, 2016; Sawtelle et al., 2012). Such differences may come from general messages about the discipline. In physics, and a few other fields, success is often viewed as a result of brilliance (Leslie et al., 2015) and women may receive fewer messages that they are brilliant and can thus succeed in physics. Such differences may also come from differential experience. In the US, women make up less than a third of students who take advanced (Physics 2 or AP Physics C) high school physics (Porter, 2019).

RQ4: Do any of the mindset factors from RQ1 predict course grade?

Despite having only small differences in SAT Math scores and compensatory strengths in HS GPA, women had lower grades in this physics course. Mindset differences, especially related to My Ability, offer a partial explanation for this phenomenon. Based on our regression models, My Growth, Others' Ability and Others' Growth did not predict course grade, while both pre- and average-My Ability did. Note, however, that less than half the grade gender/sex difference was explained by the My Ability component. It may be that other motivational factors, such as self-efficacy (Cavallo et al., 2004; Marshman et al., 2017, 2018; Nissen & Shemwell, 2016; Raelin et al., 2014; Sawtelle et al., 2012), were also important contributors to students' final grades. Alternatively, differences in the learning environment, such as micro-aggressions by peers, TAs, and instructors, or differential levels of support, may also have played an important role in the differential learning outcomes (Bian et al., 2017; Moss-Racusin et al., 2012).

Because physics self-mindset is a predictor of Physics 1 grade, finding a way to increase My Ability beliefs may mitigate gendered grade differences. In this population (primarily engineering students) women are more likely to leave the major due to concerns about low grades than men are, even when they have an A or B average (Goodman, 2002), so enhancing women's My Ability beliefs may increase retention. Importantly, average-My Ability is a stronger predictor of course grade than pre-My Ability. Thus, educators have an opportunity to intervene and

potentially improve grades and cultivate growth mindsets, especially since (from RQ1) mindset self-theory appears to be malleable during this time period. If self-mindset is simultaneously more malleable and has a stronger correlation to learning outcomes, than mindset interventions in this context should focus on students' individual experiences or the experiences of people they can relate to (for example: Binning et al., 2020; Mueller & Dweck, 1998; Walton & Cohen, 2011), rather than activities that focus on teaching students about the brain's general ability to change and grow (for example: Blackwell et al., 2007; Good et al., 2003). The latter approach appears to be well-suited to students who hold a general fixed mindset. However, it may not be useful to students who endorse a general growth mindset but a fixed self-theory. In addition to showing students that changing one's intelligence is possible, we must show them that they can change their own intelligence.

Teaching implications

For instructors who want to help students abandon fixed mindsets, student-level interventions can be valuable. It is especially important in disciplines like physics, where endorsing a fixed mindset is common (Leslie et al., 2015), that instructors clearly state that hard work and effort are necessary for success, not innate ability. Providing opportunities for self-reflection about times that students improved their abilities, or sharing stories of a diverse (so that all students in the class will be able to relate to some examples) range of people that overcame academic challenges may also help students develop a growth mindset and improve academic outcomes (Bandura, 2012; Binning et al., 2020).

Instructor-facing interventions can be useful, too. Discipline-wide mindset beliefs can predict the diversity of graduate programs (Leslie et al., 2015), but do not predict student course achievement as well as the mindset of instructors do (Canning et al., 2019). Instructors with fixed mindsets tend to have low expectations of students they believe lack natural talent, which can lead instructors to give easier assignments or encourage students to drop difficult classes because of presumed low ability (Rattan et al., 2012). Instructors with growth mindsets encourage students to accept mistakes and failures as a part of a normal learning process, congratulate persistence, and praise effort rather than intelligence when students succeed (Mueller & Dweck, 1998; Rattan et al., 2012). Instructors with growth mindsets are also more likely to implement active learning in their courses (Yik et al., 2022). Students report decreased interest in courses, as well as more concerns over fair treatment and low grades if their instructor had a fixed mindset as

opposed to a growth mindset, and this effect was larger for women than men in the study (LaCosse et al., 2021).

Limitations and future directions

The primary goal of this paper was to identify which physics intelligence mindsets participate in important empirical phenomena: changing after instruction, predictive of course grades, and potentially explaining gender/sex differences in course grades. However, it is important to acknowledge that the analyses were fundamentally correlational in nature. The causal relationship of physics intelligence mindsets would need to be further supported through intervention studies. The established benefits of other mindset interventions (e.g., Felder et al., 1995) suggest such a causal link is plausible. Further, the more specific physics intelligence mindset factor most directly associated with course grades (My Ability) suggests a new focus for mindset interventions that could have even larger effects.

A second set of concerns relate to generalizability of the findings. Because the studied institution is predominantly white, we were unable to study if mindset beliefs differ or predict grades differently for students of different racial/ethnic backgrounds due to low sample size. Although the findings were stable across the instructors in the study, a broader set of instructional contexts should also be examined. It may be that other instructional formats (e.g., with well-supported group-work) or more gender-balanced courses would produce smaller declines in physics intelligence mindsets (Haak et al., 2011). However, due to regular replications of related research (Cavallo et al., 2004; Marshman et al., 2018; Nissen & Shemwell, 2016), we believe our results are likely to translate directly to other large state universities. Results from different contexts, like such as liberal arts and community colleges, as well as schools that are much more or less selective than our institution, should be examined.

Due to the focus on gender in this study, future research should also explicitly include students who fall outside of the binary gender/sex categories included here, as well as transgender students who may not have their gender accurately recorded by the university. Though this university recently began to include more sex/gender options for students, qualitative studies may be more appropriate to understand mindset in these marginalized populations until student samples are large enough to be meaningful in quantitative analysis.

Another dimension of generalization relates to other disciplines. This study focused on gender/sex and physics mindsets because women are an underrepresented group in physics. Because the intelligence mindsets are likely important in other STEM disciplines, generalizability should be tested in other fields, especially where

women are more equitably represented (e.g., biology and chemistry). The patterns across disciplines will provide important clues into the mechanisms that produces these effects.

Conclusions

Mindset research has recently garnered attention in the physics context. This study shows that intelligence mindset can be divided into four factors: My Ability, My Growth, Others' Ability, and Others' Growth. Previous work studying mindset has divided along either by growth/ability or me/others categories, but rarely simultaneously. However, qualitative studies in physics have called for a more nuanced measurement of mindset than most surveys allow; these four categories are a step in that direction. Next, this work reveals that gender/sex differences are more pronounced in the "My" categories than the "Others" categories, and these differences are developed or exacerbated from the start to the end of an introductory physics course. These results show that women's and men's intelligence mindsets are affected differently by the classroom environment, and future studies may find this useful when developing new interventions or teaching methods aimed at helping students develop growth mindsets. Finally, we find that My Ability is the only mindset factor that predicts course grade. This information may be useful to target mindset interventions to student beliefs. A student who believes nobody can become more intelligent through hard work has very different needs than one who believes that most people can become more intelligent but that they personally lack the ability to do so.

Abbreviations

ACT: American College Testing; CFA: Confirmatory factor analysis; CFI: Comparative fit index; GRE: Graduate Record Examinations; HS GPA: High school grade point average; M: Mean; MA: My ability; MG: My growth; OA: Others' ability; OG: Others' growth; RMSEA: Root mean square error of approximation; SAT: Scholastic Achievement Test; SD: Standard deviation; SRMR: Standardized root mean square residual; STEM: Science, technology, engineering, and mathematics; TLI: Tucker–Lewis index; VIF: Variance inflation factor.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40594-022-00342-2>.

Additional file 1: Appendix A. Original Survey Items. Italicized items were removed during validation.

Additional file 2: Appendix B. Comparison of sample size, gender/sex distribution, prior academic preparation and grades between all students enrolled in the course, students who only took the pre-survey, and students who took both the pre- and post-surveys.

Additional file 3: Appendix C. To determine if men and women could be included in the same factor analysis, we conducted tests for both strong and weak measurement invariance.

Acknowledgements

The authors would like to thank Sonja Cwik, Danny Doucette, Yasemin Kalender, Yangqiuting Li, Emily Marshman, Timothy Nokes-Malach, and Kyle Whitcomb for their efforts developing the survey. We would also like to thank Robert Devaty for his manuscript feedback. Finally, we would like to thank all the students who participated in this study.

Authors' contributions

AM contributed to analysis and interpretation of data, as well as writing and revision of manuscript. CDS contributed to the interpretation of data, as well as writing and revision of manuscript. CS contributed to the conception and design of research, acquisition and interpretation of data, and revision of the manuscript. All authors read and approved the final manuscript.

Funding

This research was funded by the National Science Foundation Awards DUE-1524575 and PHY-1806691.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, USA. ²Learning, Research, and Development Center, University of Pittsburgh, Pittsburgh, PA 15260, USA.

Received: 24 August 2021 Accepted: 25 February 2022

Published online: 19 March 2022

References

- Anonson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, *38*(2), 113–125. <https://doi.org/10.1006/jesp.2001.1491>
- Bagès, C., Verniers, C., & Martinot, D. (2016). Virtues of a hardworking role model to improve girls' mathematics performance. *Psychology of Women Quarterly*, *40*(1), 55–64. <https://doi.org/10.1177/0361684315608842>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.
- Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management*, *38*(1), 9–44. <https://doi.org/10.1177/0149206311410606>
- Bian, L., Leslie, S. J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, *355*(6323), 389–391. <https://doi.org/10.1126/science.aah6524>
- Binning, K., Kaufmann, N., McGreevy, E., Fotuhi, O., Chen, S., Marshman, E., Kalender, Z., Limeri, L., Betancur, L., & Singh, C. (2020). Changing social contexts to foster equity in college science courses: An ecological-belonging intervention. *Psychological Science*, *31*(9), 1059–1070. <https://doi.org/10.1177/0956797620929984>
- Blackwell, L., Trzesniewski, K., & Dweck, C. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246–263. <https://doi.org/10.1111/j.1467-8624.2007.00995.x>
- Bleeker, M., & Jacobs, J. (2004). Achievement in math and science: Do mothers' beliefs matter 12 years later? *Journal of Educational Psychology*, *96*(1), 97–109. <https://doi.org/10.1037/0022-0663.96.1.97>
- Bouffard-Bouchard, T., Parent, S., & Larivée, S. (1991). Influence of self-efficacy on self-regulation and performance among junior and senior high-school age children. *International Journal of Behavioral Development*, *14*(2), 153.
- Browne, M., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*(2), 230–258. <https://doi.org/10.1177/0049124192021002005>

- Canning, E., Muenks, K., Green, D., & Murphy, M. (2019). STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. *Science Advances*. <https://doi.org/10.1126/sciadv.aau4734>
- Cavallo, A., Potter, W., & Rozman, M. (2004). Gender differences in learning constructs, shifts in learning constructs, and their relationship to course achievement in a structured inquiry, yearlong college physics course for life science majors. *School Science and Mathematics*, 104(6), 288–300. <https://doi.org/10.1111/j.1949-8594.2004.tb18000.x>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.
- Cook, D. A., Castillo, R. M., Gas, B., & Artino, A. R., Jr. (2017). Measuring achievement goal motivation, mindsets and cognitive load: Validation of three instruments' scores. *Medical Education*, 51(10), 1061–1074. <https://doi.org/10.1111/medu.13405>
- Crisp, G., Nora, A., & Taggart, A. (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: An analysis of students attending a Hispanic serving institution. *American Educational Research Journal*, 46(4), 924–942. <https://doi.org/10.3102/0002831209349460>
- Cwik, S., & Singh, C. (2021). How perception of learning environment predicts male and female students' grades and motivational outcomes in algebra-based introductory physics courses. *Physical Review Physics Education Research*, 17(2), 020143. <https://doi.org/10.1103/PhysRevPhysEducRes.17.020143>
- De Castella, K., & Byrne, D. (2015). My intelligence may be more malleable than yours: The revised implicit theories of intelligence (self-theory) scale is a better predictor of achievement, motivation, and student disengagement. *European Journal of Psychology of Education*, 30(3), 245–267. <https://doi.org/10.1007/s10212-015-0244-y>
- Deiglmayr, A., Stern, E., & Schubert, R. (2019). Beliefs in "brilliance" and belonging uncertainty in male and female STEM students. *Frontiers in Psychology*, 10, 1114. <https://doi.org/10.3389/fpsyg.2019.01114>
- Denworth, L. (2019, August). Debate Arises over Teaching "Growth Mindsets" to Motivate Students. *Scientific American*. Retrieved from <https://www.scientificamerican.com/article/debate-arises-over-teaching-growth-mindsets-to-motivate-students/#>
- Dweck, C. S. (2008). *Mindsets and math/science achievement*. Carnegie Corporation of New York. Retrieved from <https://www.growthmindsetmaths.com/>
- Dweck, C. (2006). *Mindset: The New Psychology of Success* (1st ed.). Random House.
- Dweck, C. (2007). Is math a gift? Beliefs that put females at risk. In S. J. Ceci & W. M. Williams (Eds.), *Why aren't more women in science? Top researchers debate the evidence* (pp. 47–55). American Psychological Association.
- Eaton, A., Saunders, J., Jacobson, R., & West, K. (2020). How gender and race stereotypes impact the advancement of scholars in STEM: Professors' biased evaluations of physics and biology post-doctoral candidates. *Sex Roles*, 82(3–4), 127–141. <https://doi.org/10.1007/s11199-019-01052-w>
- Felder, R., Felder, G., Mauney, M., Hamrin, C., Jr., & Dietz, E. (1995). A longitudinal study of engineering student performance and retention III Gender differences in student performance and attitudes. *Journal of Engineering Education*, 84(2), 151–163. <https://doi.org/10.1002/j.2168-9830.1995.tb00162.x>
- Frey, B. (2018). *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*. SAGE Publications.
- Galla, B., Shulman, E., Plummer, B., Gardner, M., Hutt, S., Goyer, J., D'Mello, S., Finn, A., & Duckworth, A. (2019). Why high school grades are better predictors of on-time college graduation than are admissions test scores: The roles of self-regulation and cognitive ability. *American Educational Research Journal*, 56(6), 2077–2115. <https://doi.org/10.3102/0002831219843292>
- Ganley, C., George, C., Cimpian, J., & Makowski, M. (2018). Gender equity in college majors: Looking beyond the STEM/non-STEM dichotomy for answers regarding female participation. *American Educational Research Journal*, 55(3), 453–487. <https://doi.org/10.3102/0002831217740221>
- García-Cepero, M. C., & McCoach, D. B. (2009). Educators' implicit theories of intelligence and beliefs about the identification of gifted students. *Universitas Psychologica*, 8(2), 295–310.
- Gonzalez, H. B., & Kuenzi, J. J. (2012). *Science, Technology, Engineering, and Mathematics (STEM) Education: A Primer* (CRS Report No. R42530). Library of Congress Congressional Research Service.
- Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, 24(6), 645–662. <https://doi.org/10.1016/j.appdev.2003.09.002>
- Goodman, I. (2002). *Final report of the women's experiences in college engineering project* (ED507394). ERIC. Retrieved from <https://files.eric.ed.gov/fulltext/ED507394.pdf>
- Haak, D., Hillerislanders, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science*, 332(6034), 1213–1216. <https://doi.org/10.1126/science.1204820>
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.
- Hazari, Z., Sonnert, G., Sadler, P., & Shanahan, M. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching*, 47(8), 978–1003. <https://doi.org/10.1002/tea.20363>
- Hazari, Z., Tai, R., & Sadler, P. (2007). Gender differences in introductory university physics performance: The influence of high school physics preparation and affective factors. *Science Education*, 91(6), 847–876. <https://doi.org/10.1002/sce.20223>
- Hu, L., & Bentler, P. (1999). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424–453. <https://doi.org/10.1037/1082-989X.3.4.424>
- Integrated Postsecondary Education Data System (IPEDS). (2020). *University of Pittsburgh-Pittsburgh Campus*. U.S. Department of Education, National Center for Education Statistics. Retrieved from <https://nces.ed.gov/ipeds/datacenter/institutionprofile.aspx?unitid=215293>
- Jacobs, J., & Eccles, J. (1992). The impact of mothers' gender-role stereotypic beliefs on mothers' and children's ability perceptions. *Journal of Personality and Social Psychology*, 63(6), 932–944. <https://doi.org/10.1037/0022-3514.63.6.932>
- Kalender, Y. Z., Marshman, E., Schunn, C. D., Nokes-Malach, T. J., & Singh, C. (2022). Framework for unpacking gendered mindsets in physics by students' gender. *Physical Review Physics Education Research*, 18, 010116. <https://doi.org/10.1103/PhysRevPhysEducRes.18.010116>
- Kepple, C., Tilahun, M., Matti, N., & Coble, K. (2020). Pedagogy training for the development of GTA mindsets and inclusive teaching practices. In S. Wolf, M. B. Bennett, & B. W. Frank (Eds.), *Proceedings of the 2020 Physics Education Research Conference* (pp. 272–277). American Association of Physics Teachers. <https://doi.org/10.1119/perc.2020.pr.Kepple>
- Kline, R. B. (2016). *Methodology in the social sciences. Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Kost, L., Pollock, S., & Finkelstein, N. (2009). Characterizing the gender gap in introductory physics. *Physical Review Physics Education Research*, 5(1), 010101. <https://doi.org/10.1103/PhysRevSTPER.5.010101>
- Kricorian, K., Seu, M., Lopez, D., Ureta, E., & Equils, O. (2020). Factors influencing participation of underrepresented students in STEM fields: Matched mentors and mindsets. *International Journal of STEM Education*, 7(1), 1–9. <https://doi.org/10.1186/s40594-020-00219-2>
- LaCosse, J., Murphy, M. C., Garcia, J. A., & Zirkel, S. (2021). The role of STEM professors' mindset beliefs on students' anticipated psychological experiences and course interest. *Journal of Educational Psychology*, 113(5), 949–971. <https://doi.org/10.1037/edu0000620>
- Leslie, S., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, 347(6219), 262–265. <https://doi.org/10.1126/science.1261375>
- Limeri, L., Carter, N., Choe, J., Harper, H., Martin, H., Benton, A., & Dolan, E. (2020). Growing a growth mindset: Characterizing how and why undergraduate students' mindsets change. *International Journal of STEM Education*, 7(1), 1–19. <https://doi.org/10.1186/s40594-020-00227-2>
- Little, A., Humphrey, B., Green, A., Nair, A., & Sawtelle, V. (2019). Exploring mindset's applicability to students' experiences with challenge in transformed college physics courses. *Physical Review Physics Education Research*, 15(1), 010127. <https://doi.org/10.1103/PhysRevPhysEducRes.15.010127>
- Madsen, A., McKagan, S., & Sayre, E. (2013). Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap? *Physical Review Physics Education Research*, 9(2), 020121. <https://doi.org/10.1103/PhysRevSTPER.9.020121>

- Marshman, E., Kalender, Z., Nokes-Malach, T., Schunn, C., & 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? *Physical Review Physics Education Research*, 14(2), 020123. <https://doi.org/10.1103/PhysRevPhysEducRes.14.020123>
- Marshman, E., Kalender, Z., Schunn, C., Nokes-Malach, T., & Singh, C. (2017). A longitudinal analysis of students' motivational characteristics in introductory physics courses: Gender differences. *Canadian Journal of Physics*, 96(4), 391–405. <https://doi.org/10.1139/cjp-2017-0185>
- Miller, K., Schell, J., Ho, A., Lukoff, B., & Mazur, E. (2015). Response switching and self-efficacy in Peer Instruction classrooms. *Physical Review Physics Education Research*, 11(1), 010104. <https://doi.org/10.1103/PhysRevSTPER.11.010104>
- Moss-Racusin, C., Dovidio, J., Brescoll, V., Graham, M., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, 109, 16474.
- Mueller, C., & Dweck, C. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, 75(1), 33–52. <https://doi.org/10.1037/0022-3514.75.1.33>
- Muenks, K., & Miele, D. B. (2017). Students' thinking about effort and ability: The role of developmental, contextual, and individual difference factors. *Review of Educational Research*, 87(4), 707–735. <https://doi.org/10.3102/0034654316689328>
- National Center for Science and Engineering Statistics. (2019). *Women, minorities, and persons with disabilities in science and engineering: 2019* (Special Report NSF 19-304). National Science Foundation. Retrieved from <https://www.nsf.gov/statistics/wmpd>.
- Nissen, J., & Shemwell, J. (2016). Gender, experience, and self-efficacy in introductory physics. *Physical Review Physics Education Research*, 12(2), 020105. <https://doi.org/10.1103/PhysRevPhysEducRes.12.020105>
- Pintrich, P., & De Groot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33–40. <https://doi.org/10.1037/0022-0663.82.1.33>
- Porter, A. (2019). *Women in Physics and Astronomy 2019*. Statistical Research Center of the American Institute of Physics. Retrieved from <https://www.aip.org/sites/default/files/statistics/women/Women%20in%20Physics%20and%20Astronomy%202019.1.pdf>
- Raelin, J., Bailey, M., Hamann, J., Pendleton, L., Reisberg, R., & Whitman, D. (2014). The gendered effect of cooperative education, contextual support, and self-efficacy on undergraduate retention. *Journal of Engineering Education*, 103(4), 599–624. <https://doi.org/10.1002/jee.20060>
- Rattan, A., Good, C., & Dweck, C. (2012). "It's ok—Not everyone can be good at math": Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology*, 48(3), 731–737. <https://doi.org/10.1016/j.jesp.2011.12.012>
- Rattan, A., Savani, K., Chugh, D., & Dweck, C. S. (2015). Leveraging mindsets to promote academic achievement: Policy recommendations. *Perspectives in Psychological Science*, 10, 721–726. <https://doi.org/10.1177/1745691615599383>
- Rattan, A., Savani, K., Komarraju, M., Morrison, M. M., Boggs, C., & Ambady, N. (2018). Meta-lay theories of scientific potential drive underrepresented students' sense of belonging to science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology*, 115(1), 54–75. <https://doi.org/10.1037/pspi0000130>
- Redish, E., Saul, J., & Steinberg, R. (1998). Student expectations in introductory physics. *American Journal of Physics*, 66(3), 212–224. <https://doi.org/10.1119/1.18847>
- Sadler, P., & Tai, R. (2001). Success in introductory college physics: The role of high school preparation. *Science Education*, 85(2), 111–136. [https://doi.org/10.1002/1098-237X\(200103\)85:2%3C111::AID-SCE20%3E3.0.CO;2-O](https://doi.org/10.1002/1098-237X(200103)85:2%3C111::AID-SCE20%3E3.0.CO;2-O)
- Salehi, S., Burkholder, E., Lepage, G., Pollock, S., & Wieman, C. (2019). Demographic gaps or preparation gaps? The large impact of incoming preparation on performance of students in introductory physics. *Physical Review Physics Education Research*, 15(2), 020114. <https://doi.org/10.1103/PhysRevPhysEducRes.15.020114>
- Sawtelle, V., Brewes, E., & Kramer, L. (2012). Exploring the relationship between self-efficacy and retention in introductory physics. *Journal of Research in Science Teaching*, 49(9), 1096–1121. <https://doi.org/10.1002/tea.21050>
- Scherr, R., Plisch, M., Gray, K., Potvin, G., & Hodapp, T. (2017). Fixed and growth mindsets in physics graduate admissions. *Physical Review Physics Education Research*, 13(2), 020133. <https://doi.org/10.1103/PhysRevPhysEducRes.13.020133>
- Schudson, Z. C. (2021). Psychology's stewardship of gender/sex. *Perspectives on Psychological Science*, 16(6), 1105–1112. <https://doi.org/10.1177/17456916211018462>
- Shih, S. (2011). Perfectionism, implicit theories of intelligence, and Taiwanese eighth-grade students' academic engagement. *The Journal of Educational Research*, 104(2), 131–142. <https://doi.org/10.1080/00220670903570368>
- Sisk, V., Burgoyne, A., Sun, J., Butler, J., & Macnamara, B. (2018). To what extent and under which circumstances are growth mind-sets important to academic achievement? Two meta-analyses. *Psychological Science*, 29(4), 549–571. <https://doi.org/10.1177/0956797617739704>
- StataCorp. (2021). *Stata statistical software: Release 17*. StataCorp LLC.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797–811. <https://doi.org/10.1037/0022-3514.69.5.797>
- The College Board and ACT, Inc. (2018). *Guide to the 2018 ACT/SAT concordance*. Retrieved from <https://www.act.org/content/dam/act/unsecured/documents/ACT-SAT-Concordance-Information.pdf>
- Traxler, A., Cid, X., Blue, J., & Barthelemy, R. (2016). Enriching gender in physics education research: A binary past and a complex future. *Physical Review Physics Education Research*, 12, 020114. <https://doi.org/10.1103/PhysRevPhysEducRes.12.020114>
- Troche, S. J., & Kunz, A. (2020). The factorial structure and construct validity of a German translation of Dweck's Implicit Theories of Intelligence Scale under consideration of the wording effect. *Psychological Test and Assessment Modeling*, 62(3), 386–403.
- University of Pittsburgh. (2021). *Student diversity dashboard*. Retrieved from <https://www.diversity.pitt.edu/social-justice/student-dashboard>
- Van Dusen, B., & Nissen, J. (2020). Associations between learning assistants, passing introductory physics, and equity: A quantitative critical race theory investigation. *Physical Review Physics Education Research*, 16, 010117. <https://doi.org/10.1103/PhysRevPhysEducRes.16.010117>
- Vincent-Ruz, P., Binning, K., Schunn, C., & Grabowski, J. (2018). The effect of math SAT on women's chemistry competency beliefs. *Chemistry Education Research and Practice*, 19(1), 342–351. <https://doi.org/10.1039/C7RP00137A>
- Vooren, M., Haelermans, C., Groot, W., & van den Brink, H. M. (2022). Comparing success of female students to their male counterparts in the STEM fields: an empirical analysis from enrollment until graduation using longitudinal register data. *International Journal of STEM Education*. <https://doi.org/10.1186/s40594-021-00318-8>
- Voyer, D., & Voyer, S. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140(4), 1174–1204. <https://doi.org/10.1037/a0036620>
- Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes among minority students. *Science*, 331(6023), 1447–1451. <https://doi.org/10.1126/science.1198364>
- Whitcomb, K., & Singh, C. (2020). For physics majors, gender differences in introductory physics do not inform future physics performance. *European Journal of Physics*, 41, 065701. <https://doi.org/10.1088/1361-6404/ab9f1d>
- Whitcomb, K. M., Kalender, Z. Y., Nokes-Malach, T. J., Schunn, C. D., & Singh, C. (2021). Comparison of self-efficacy and performance of engineering undergraduate women and men. *International Journal of Engineering Education*, 26(6), 1996–2014. https://www.ijee.ie/1atestissues/Vol36-6/24_ijee4004.pdf
- Yeager, D., & Dweck, C. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, 47(4), 302–314. <https://doi.org/10.1080/00461520.2012.722805>
- Yeager, D. S., & Dweck, C. S. (2020). What can be learned from growth mindset controversies? *American Psychologist*, 75(9), 1269–1284. <https://doi.org/10.1037/amp0000794>
- Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnow, R., Muller, C., Tipton, E., Schneider, B., Hinojosa, C. P., Paunesku, D., Romero, C., Flint, K., Roberts, A., Troll, J., Iachan, R., Buontempo, J., Yang, S. M., & Carvalho, C., ... Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, 573(7774), 364–369. <https://doi.org/10.1038/s41586-019-1466-y>
- Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., Lee, H. Y., O'Brien, J., Flint, K., Roberts, A., Trott, J., Greene, D., Walton, G.

M., & Dweck, C. S. (2016). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of Educational Psychology, 108*(3), 374–391. <https://doi.org/10.1037/edu0000098>

Yik, B. J., Raker, J. R., Apkarian, N., Stains, M., Henderson, C., Dancy, M. H., & Johnson, E. (2022). Evaluating the impact of malleable factors on percent time lecturing in gateway chemistry, mathematics, and physics courses. *International Journal of STEM Education*. <https://doi.org/10.1186/s40594-022-00333-3>

Zeidner, M. (1998). *Test anxiety: the state of the art*. Plenum Press.

Zimmerman, B. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*(1), 82–91. <https://doi.org/10.1006/ceps.1999.1016>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
