Received: 10 January 2019 | Revised: 18 March 2020 | Accepted: 17 April 2020

DOI: 10.1002/sce.21580

Science Education WILEY

GENERAL SECTION

The STEM grading penalty: An alternative to the "leaky pipeline" hypothesis

Dirk Witteveen[1] | Paul Attewell[2]

[1]Nuffield College, University of Oxford, Oxford, UK

[2]Sociology Department, The Graduate Center, City University of New York, New York, New York

Correspondence
Dirk Witteveen, Nuffield College, University of Oxford, 1 New Road, Oxford, OX1 1NF, United Kingdom.
Email: dirk.witteveen@nuffield.ox.ac.uk

Abstract

The low number of baccalaureates in science, technology, engineering, and mathematics (STEM) is often viewed as problematic for the US's economic competitiveness, leading scholars to search for explanations for STEM retention. Our analyses of the Beginning Postsecondary Students Longitudinal Study indicate that the notion of a so-called "leaky STEM pipeline" out of STEM majors overstates the problem because it neglects the substantial influx into STEM from other majors throughout college. Researchers concerned with STEM retention should focus on a broader defined group of "STEM-actives": A combination of freshman students who declared a STEM major or who take a considerable number of STEM credits. Among these students ($N = 3{,}020$) we examine the variation in the relatively lower grades that many individuals earn in STEM courses compared to their non-STEM courses. The size of an undergraduate's "STEM-grading penalty"—an individual grading disparity—in the first couple of college semesters is significantly associated with the probability of leaving STEM. The influence of this STEM-penalty on STEM graduation chances is robust to college students' variation in both general academic achievement and STEM-specific preparation, thereby eliminating a large portion of the effect

due to skills, performance, and selection. Our analyses expands on previous research regarding relative grading conducted *within* STEM-fields.

KEYWORDS
college major, higher education, relative grading, retention, STEM

1 | INTRODUCTION

Ever since the Sputnik crisis of 1957, educators and policymakers have argued that the United States is losing its international lead in technology due to weaknesses in the educational preparation and/or commitment of students in the sciences, technology, engineering, and mathematics—fields known collectively as STEM (Geiger, 1997). Sixty years later, this concern persists, with high-profile reports claiming that the US is failing to adequately develop its native-born talent, leading the nation to become heavily dependent on immigrants for STEM-literate workers and researchers (Carnevalle, Smith, & Melton, 2011). Prominent among these reports are the *Rising Above the Gathering Storm* by the National Academy (2007), *The Opportunity Equation* by the Carnegie Corporation and the Institute for Advanced Study (2009), and *Engage to Excel* by the Executive Office of the President (2012), which state that the US economy will be short of one million STEM workers in the next decade if the current STEM graduation rate remains the same. It also calculated that a 10%-point improvement of the STEM retention rate would cover about three-quarters of this national target. In addition, several empirical studies have shown that increases in STEM workers positively affect the growth of wages and total productivity, and have argued that the demand for STEM work is coming from both the government sector and private industries (Peri, Shih, & Sparber, 2015; Xue & Larson, 2015).

Research on STEM enrollment typically highlights the fact that students who initially had an interest are *leaking out* of STEM somewhere between K–12 education and college graduation. A third of all 4-year college freshmen express interest in STEM fields before taking any college coursework and 28% of all bachelor's (BA) degree students enter a STEM field at *some point* during their higher education career (National Science Board, 2012). Yet this broad initial interest in STEM seems to decline rapidly during later college years. A striking 48% of students who entered a BA STEM program had left these fields entirely within 6 years (Chen & Soldner, 2013). Steady attrition from STEM subjects occurs over time, leaving fewer students studying STEM as one moves toward the end of the STEM pathway (Goulden, Frasch, & Mason, 2009; Kokkelenberg & Sinha, 2010; Maltese & Tai, 2011; Nicholls, Wolfe, Besterfield-Sacre, Shuman, & Larpkiattaworn, 2007). Researchers and policymakers have labeled this phenomenon the "leaky STEM pipeline" (Blickenstaff, 2005; Metcalf, 2010).

However, other researchers have criticized the leaky pipeline imagery (Cannady, Greenwald, & Harris, 2014; Metcalf, 2010; Xie & Killewald, 2012). While acknowledging processes that progressively narrow the numbers leading to STEM graduation, they argue that the leaky pipeline imagery implies that there is a single normative route to a STEM degree, when in fact there are multiple routes. The leaky pipeline metaphor also suggests that "patching the leaks" would be sufficient to increase STEM graduation, a policy implication that some scholars have questioned (Cannady et al., 2014, p. 444). Others have shown that STEM graduates have rather varied course-taking and motivational histories, whereby their final interest in STEM and pursuing a STEM degree varied across a wide age-span (Maltese, Melki, & Wiebke, 2014).

In this study, we compare the STEM pipeline approach to the alternative route approach. Using longitudinal data for BA degree attendees, we first present an empirically driven argument demonstrating that the leaky pipeline is an unhelpful metaphor for understanding the selection and sorting mechanisms within STEM higher

education. Our analysis of first-year students' course-taking behavior and major selection indicates that the STEM attrition is about as large as the *influx* into STEM and that attrition from traditional STEM majors is part of a general attrition dynamic in higher education which is not specific to STEM majors. We also suggest a research- and policy-focus on a broader group of "STEM-active students."

Subsequently, we address two substantive questions regarding mechanisms of STEM attrition among the identified "STEM-active" student population. (a) Is there a substantive grading "penalty"—a lower STEM grade point average (GPA) than non-STEM GPA—among STEM-active college students? (b) Is relative grading negatively associated with individuals' likelihood of persisting in STEM while controlling for academic and STEM-performance measures?

The "STEM-grade penalty," as discussed in this paper, consists of a student receiving lower grades in STEM courses during early semesters in college, compared to that same student's grades in other types of classes during the same semesters. Further analyses indicate that even students who persist—and ultimately do graduate in STEM—weather several semesters of relatively low grades in their STEM courses, whereas other (more) penalized undergraduates retreat from STEM into other disciplines or stop out of college without completing a degree. This suggests that STEM attrition in college should partly be understood as a student response to disincentives created by lower *relative* course grades or as a response to a "message" implied by grades about where a student's true academic talents lie.

2 | FRAMEWORK

The research questions and analyses rely on two perspectives in educational research. We first address the definition of STEM students to refine the applicability of the "leaky STEM pipeline" approach to STEM attrition (Blickenstaff, 2005; Metcalf, 2010). We make an important analytical decision by empirically defining a STEM student at the beginning of the college pipeline—one based on behavior regarding course-taking rather than simply pursuing or declaring one of the STEM majors at college entry. Importantly, this definitional choice allows for the possibility that *non*-STEM students who take considerable STEM coursework in their first college semester either continue to do so and organically become STEM students or change their minds and switch between majors (to STEM). These decisions can be understood as "rational" because, as students make progress in college, they continue to update information about academic subjects, their achievements, and their preferences or interests (Altonji, Blom, & Meghir, 2012; Maltese et al., 2014; Nauta, 2007). As a result, in this study, the exodus from STEM is studied while taking into account the often-unobserved influx of STEM students coming into the STEM "pipeline."

The second question addresses the role of grading for STEM major persistence. Here, we draw on existing knowledge regarding *relative* grading specific in STEM as discussed by Ost (2010), Rask (2010), and others, as discussed later. Regardless of discipline or subject, grades have a substantial impact of subsequent course selection because they operate as signals to students which may push them away from certain subjects or pull them toward others, as has been shown in numerous education studies (e.g. Bar, Kadiyali, & Zussman, 2009; Fournier & Sass, 2000). We regard the relative grading hypothesis as the best-suited analytical lens for understanding how performance signals contribute to moving away from STEM education, especially among a subsample consisting of academically strong (i.e., graduating) and STEM-interested or STEM-involved students. Moreover, these students often have shown above-average achievement in math during K–12 schooling (Maltese & Tai, 2011; Nicholls et al., 2007). Thus far, the relative grading framework has been used to describe the grading discrepancies between academic departments in US colleges. In this study, we estimate the association between relative grading and leaving STEM by directly observing students' grades in STEM and non-STEM coursework throughout several consecutive semesters.

3 | LITERATURE AND HYPOTHESES

3.1 | Path dependency in the STEM pipeline

Explanations for STEM attrition already prominent in the literature include studies of students' demographic characteristics, educational preparation, previous course trajectories, college momentum and performance, and competition with other majors (Xie, Fang, & Shauman, 2015). There is strong evidence among these studies for the importance of academic preparation, before college entry, in explaining why the STEM pipeline is "leaky" (Chen & Soldner, 2013). For example, poor average grades in high school have been shown to be predictive of STEM attrition in college (Kokkelenberg & Sinha, 2010; Shaw & Barbuti, 2010; Whalen & Shelley, 2010). Furthermore, elementary math courses in high school serve as "gatekeepers" for subsequent acquisition of math skills (Long, Conger, & Iatarola, 2012) and, eventually, college attendance (Aughinbaugh, 2012).

Most studies, however, concentrate on the reverse perspective—the determinants of STEM completion. Research has shown the importance of science courses and the level of high school math completed (Federman, 2007; Griffith, 2010; Harackiewicz, Barron, Tauer, & Elliot, 2002; Redmond-Sanogo, Angle, & Davis, 2016; Riegle-Crumb & King, 2010; Sadler, Sonnert, Hazari, & Tai, 2014; Tyson, Lee, Borman, & Hanson, 2007), with particular emphases on the positive role of middle-school algebra and high school calculus (Maltese & Tai, 2011; Nicholls et al., 2007; Sadler et al., 2014) and the number of advanced placement (AP) courses taken (Robinson, 2003). In addition, having an expressed ambition for a STEM degree or a STEM profession seems to have a positive influence on persistence in the pipeline, often measured by declaring an interest in STEM careers or majors early on (Daempfle, 2002; Rask, 2010; Tai, Liu, Maltese, & Fan, 2006; Tyson, 2011), and tied to occupational interests (Morgan, Gelbgiser, & Weeden, 2013).

However, scholars who have reservations about the leaky STEM pipeline approach typically stress its over-concentration on the single modal pathway, in particular, in relation to academic preparation and taking math. For instance, high school students are not *necessarily* leaving the STEM pipeline if they do not fulfill coursework benchmarks (grades). Having motivated and talented peers is equally important (Ost, 2010). In addition, Cannady et al. (2014) stressed the image of a "multilayered STEM pipeline," primarily because ambitious students in other college majors, with similar math backgrounds, could still *enter* a STEM track later on.

3.2 | Gender and race

Scholars found that women, underrepresented minorities, first-generation college students, and students from low-income families are more likely to leave STEM than their counterparts, as summarized in a review by Chen and Soldner (2013). In explaining the gender gap in STEM graduation, educational researchers have identified hurdles and disadvantages at various steps in the pipeline. Some studies point to cultural socialization throughout childhood; the broader societal context and influence of gender bias and gender stereotypes (Blickenstaff, 2005; Hill, Corbett, & St. Rose, 2010). These include gender stereotypes of scientists in public media or textbooks (Stout, Dasgupta, Hunsinger, & McManus, 2011), leading to gender-specific expectations regarding occupational preferences during secondary education (Ceci, Williams, & Barnett, 2009). As a result, girls in high school display less interest and confidence in math (Catsambis, 1994) and a lower level of perceived competence net of performance (Correll, 2001).

These gendered STEM disadvantages typically extend into higher education, where men disproportionally enroll in STEM majors, in particular in physical science and engineering (Riegle-Crumb & King, 2010). After accounting for prior levels of achievement, women remain less likely to enroll in a STEM major at college entry. This led some scholars to suggest that part of the "unexplained variance" remains attributable to gender-specific preferences and purposeful choices (Diekman, Brown, Johnston, & Clark, 2010; Riegle-Crumb, King, Grodsky, & Muller, 2012). The gender gap in STEM attainment has been decreasing at the BA level during the past decade (Miller & Wai, 2015; Simpson, 2001).

Among students who have entered a STEM major, women are also less likely to persist (Alper & Gibbons, 1993), regardless of their actual educational performance (Fischer, 2017). Furthermore, women have become less likely to pursue a PhD after a STEM BA degree in recent years (Miller & Wai, 2015). Aside from the argument that women have different preferences, scholars have shown that a substantial proportion of the gender gap in STEM-major persistence is rooted in differences in preparation (Griffith, 2010). However, a substantial body of research points to the chilly climate for women in STEM programs as an explanation for the gender gap in STEM retention (Diekman et al., 2010; Marra, Rodgers, Shen, & Bogue, 2012; Ong, Wright, Espinosa, & Orfield, 2011; Seymour & Hewitt, 1997), such as unfavorable interactions between women STEM students and their college advisors (Gayles & Ampaw, 2014) or with high-performing (male) peers (Fischer, 2017). Moreover, focusing on contexts that contribute to women's persistence, researchers have emphasized the positive influences of female instructors (Bottia, Stearns, Mickelson, Moller, & Valentino, 2015) and of a high proportion of female graduate students in the same institution (Griffith, 2010; see Wang & Degol, 2013 for a comprehensive review of gender differences in STEM).

With regard to the race and ethnicity gap in STEM, Black and Hispanic students do not have lower levels of interest in science fields (Simpson, 2001; Xie et al., 2015). Instead, the underrepresentation of Black and Hispanic students in STEM is partially rooted in their underrepresentation in college-prep courses and their lower grades in high school science (Griffith, 2010; Hurtado et al., 2007). Moreover, high levels of college debt among Hispanic students may negatively affect their decisions to enroll and persist in STEM (Malcom & Dowd, 2008), although Crisp, Nora, and Taggart (2009) have disputed this mechanism. Comparable to positive factors for women's STEM persistence, Black students tend to be more likely to persist in a STEM major if their classes are taught by a Black instructor (Price, 2010).

3.3 | College course-taking and advisement

In addition to academic preparation before entering higher education, researchers have found that math performance *in college* can be linked to the choice of major and to STEM students leaving a STEM program. Few students who are required to take remedial math courses in college complete the remedial sequence or pass a "college-level" math course (Jaggars & Hodara, 2011) or pursue STEM majors later on (Bettinger, 2010). Bettinger (2010) also found that students in Ohio who took a high proportion of STEM courses in their first semester of college had better STEM retention records than otherwise similar counterparts who took fewer STEM courses as part of a more varied curriculum. The PCAST (2012) report suggested that some students who want a STEM degree are held back by their inadequate math skills—the initial reason for remediation in college—*as well as* by the actual remediation that follows.

National surveys of student engagement have examined pedagogical practices as reported by STEM and non-STEM senior undergraduates. Though STEM seniors are more likely than their non-STEM peers to undertake collaborative learning, "Seniors majoring in arts and humanities observed the highest levels of effective teaching practices, whereas those in STEM fields, especially engineering, observed the lowest levels" (National Survey of Student Engagement, 2013, p. 8). Those effective practices included professors' clarity and organization in lectures, their use of examples, and the provision of formative feedback on assignments and finals. Recent educational research has, therefore, focused on policies to improve the quality of STEM pedagogy (Fairweather, 2009; Gehrke & Kezar, 2017).

3.4 | College performance and grading

College science courses have a reputation among students for being difficult, leaving students vulnerable to receiving much lower grades. For an early generation of undergraduates, Becker, Geer, and Hughes (1968)

discovered that students were very concerned about "making the grade"—a preoccupation with maintaining a high GPA. For the present generation of college students, this might be read as students' concern with protecting or maximizing their GPA. Students perceive good grades as important both as a recognition of their skills and effort and view bad grades as impairing their eligibility for merit-based aid, or their chances of graduating *cum laude*, or their likelihood of admission into higher degree programs after graduation. The literature on grading has reported substantial empirical evidence for the influence of one's performance in introductory courses on final undergraduate major *choice* (e.g., Chizmar, 2000; Main, Mumford, & Ohland, 2015; Rask, 2010; Rask & Bailey, 2002; Rask & Tiefenthaler, 2008; Sabot & Wakeman-Linn, 1991).

In addition, grades received in early coursework matter for students' decision to stay in a particular college major. Rask (2010) showed that the *absolute grade* received in the first semester is the strongest predictor of persisting in that department's coursework for the following second and third semesters. This mechanism is particularly strong among STEM majors. Using different data, Main et al. (2015) confirmed this grade "pull effect" in STEM among several engineering majors. The authors also showed that this effect becomes stronger if students expected to receive higher grades in that major's upper-division courses. Thus, first-semester course performance (absolute grades) in STEM subjects may serve as "gatekeepers" for subsequent courses in STEM majors. Moreover, several researchers have speculated that receiving *low* grades in early STEM courses may be a "push factor" (away from STEM); shifting students into non-STEM subjects and ultimately into declaring a non-STEM major (Hrabowski, 2012; Ost, 2010; Seymour & Hewitt, 1997).

Aside from receiving grades received in key courses early on, students' decisions to leave STEM may be based on so-called *relative grades*—receiving a lower grade in one discipline compared to another. Initial studies on relative grading were initially concerned with the extent to which average grading levels vary across disciplines over time. This is grade inflation. Grading practices and standards differ substantially between STEM majors and non-STEM majors. In general, grades given in STEM departments are often among the lowest of college departments (see Ost, 2010; Sabot & Wakeman-Linn, 1991). More precisely, STEM course grading typically uses a much wider score distribution, making more use of the lower end of the grading scale.

This strong between-department contrast makes relative grading another potentially important factor in students' decision-making. If students directly compare their grades received in STEM and their grades received outside of STEM, this (potential) gap may form *another* matching signal for persistence or exiting a STEM major. However, thus far, the empirical evidence of relative grading affecting STEM-major persistence and graduation has not been as strong as for absolute grading. Evidence from a small liberal arts college suggested that low relative grades indeed negatively affect persistence in economics, although this was only apparent among women, but not among men (Rask & Tiefenthaler, 2008). A follow-up study on longitudinal data from the same college partially contradicted those earlier findings. It reported that the relative grade effect can indeed be observed across its institution's STEM majors, but that it varies by discipline (strongest for chemistry) and sometimes applies to the first or the second course taken. Furthermore, relative grading affects men much more strongly than women (Rask, 2010).

Together, the somewhat conflicting findings regarding relative grading in college-specific datasets leads us to a key hypothesis regarding STEM attrition across US higher education. Using nationally representative data we expect to find a negative influence of lower *relative* grades (i.e., an individual's grades in STEM courses, relative to that same individual's non-STEM course grades) on persistence in STEM. Importantly, if receiving lower STEM grades is a relevant mechanism of attrition among STEM students, its effect should be observable even among students with academically strong records and good STEM preparation: net of academic preparation and absolute grades *across* disciplines. An average shift away from STEM coursework among this high-performing group constitutes the "STEM-penalty hypothesis."

3.5 | Competition from other majors

American universities differ from those in many other countries in allowing undergraduates to choose a major after being admitted into the university. The fact that US undergraduates are allowed to choose among courses and majors while in college, generates various competitive and gatekeeping processes within colleges. Some popular majors are oversubscribed and use a higher GPA or course prerequisites to limit the numbers or to "cream" students entering that course of study. Other majors are less popular and less likely to have high barriers to entry.

One consequence of this system is that many American undergraduates try out a major before finally committing to one: approximately 28% of undergraduates in 4-year colleges declare one major and later switch (Chen & Soldner, 2013, p. 15). Stinebrickner and Stinebrickner (2014) argued that one explanation for low STEM graduation is the fact that many potentially talented science students try out (too many) courses in different disciplines. More early exposure to STEM coursework might lead to a higher STEM influx and, therefore, more STEM graduates. We will focus on the role of competition from other majors in this study.

Furthermore, the "STEM pipeline narrative" suggests that leaving STEM is more problematic and structurally different from other college disciplines. However, research on college major migration conducted in the 1990s has shown that students exiting STEM have a lot in common with students exiting non-STEM majors. Seymour and Hewitt (1997) used ethnographic methods documenting that "switchers" report remarkably similar reasons for doing so, such as a perception of lack of success and a dissatisfaction with the way courses are taught. They emphasized that the actual performance (academic quality) of switchers is the same across college majors, including STEM. More recently, Ohland et al. (2008) confirmed the consistency of major switching across disciplines by quantitatively analyzing persistence rates and engagement. They also found that students of some STEM majors, in particular, engineering, even display *higher* levels of persistence than other college majors.

4 | DATA AND ANALYTICAL APPROACH

This study applies an observational research design to panel data. The analyses presented below, draw upon a longitudinal study directed by the National Center for Education Statistics (2011) that tracks a nationally representative cohort of first-time freshmen for 6 years after their initial entry into college in 2004. Respondents in this wave of the Beginning Postsecondary Students Longitudinal Study (BPS) were interviewed towards the end of their first year of college and then 3 and 6 years after the first entry. Each student record contains information from all colleges attended during this period (Wine, Janson, & Wheeless, 2011). We analyze the Restricted-Use Transcript Data Files of the BPS to measure student progress semester by semester. We use weights developed by the BPS to adjust for attrition and nonresponse. In addition, BPS replaced missing data using hot-deck imputation. We have rounded the sample sizes reported below to the nearest 10, as required by the National Center for Education Statistics' restricted data license.

Our transcript sample contains 8,980 students who were registered in a 4-year college for their first semester in Fall 2004. A total of 5,500 of these students (61%) graduated with a BA degree within 6 years. Among those graduated, 1,000 (18%) obtained their degree in a STEM field: Biological & biomedical sciences, computer & information science, engineering, mathematics & statistics, or physical sciences. Consistent with a majority of previous academic and policy research, we exclude the applied technical or "technology-related" programs from the definition of a STEM major. Using students' transcript data, STEM courses are systematically defined by the National Center for Education Statistics (2011) using the "College Course Map." This classification provides the best possible comparability across institutions.

In addition to identifying by students' responses to the survey question ("What is your intended/declared college major?"), we evaluate the student's actual activity in STEM courses to define a broader group of "STEM-actives"

among the total student population of future bachelor recipients. Among this analytically relevant subgroup—STEM-actives—we address the following research questions.

First: What is the net attrition out of STEM? Is the exodus from traditional STEM majors counterbalanced by an influx from other majors? This question leads us to revisit the question of "who is a STEM student?"

Second: Among successful (eventual graduates) and high-performing students who enroll in STEM coursework, what are the patterns of relative grading as measured by one's STEM-GPA and non-STEM-GPA across the first semesters in college? Do these averages vary by gender?

Third: Is relative grading—a "STEM grading penalty"—negatively associated with one's chances of STEM retention over and above demographic factors, academic preparation, and early course selection?

On the basis of the literature, we expect to find evidence for a "multilayered STEM pipeline" (Cannady et al., 2014), rather than a single modal pathway, in the form of a substantive influx from non-STEM majors throughout college. Furthermore, among students who are active in STEM coursework early on, we expect to find variation in the balance between the averages of STEM grades and non-STEM grades, over time, between students who persist in STEM and those who do not. As this usage of grading variation is relatively novel to the study of STEM attrition, we remain agnostic with regard to the strength of this grading gap. Finally, we expect a negative influence of receiving lower *relative* grades (i.e., a "STEM penalty") on persistence in STEM within the observation period (6 years after college entry).

To test these hypothesized effects, we estimate logistic regression models on graduation in STEM, among a sufficiently large subsample of early STEM-active BA graduates ($N = 3{,}020$). Of critical importance is the construction of the "STEM-penalty" variable. Within our subsample of STEM course-taking students, we build a difference score by subtracting each student's non-STEM GPA from that same student's STEM GPA (for each of the first four semesters). A negative value on this variable indicates a within-student *and* within-semester grading penalty. This measure allows us to estimate the impact of relative grading in *subsequent* semesters on STEM-attrition.

Previous studies emphasized the positive role of good absolute grades for persistence in specific STEM subjects (such as economics, physics, and biology). The current study builds on previous research of relative grading. We improve on earlier research that has been conducted with data from a single US college, focusing on the course-to-course pattern (Rask, 2010; Rask & Tiefenthaler, 2008), and research conducted on specific STEM majors (e.g., Main et al., 2015). Our approach differs in several ways from these preceding studies that identified both significant and weak effects of relative grading on college major (STEM) persistence. Most importantly, we use national-representative data to test whether a general (STEM vs. non-STEM) grading *penalty* can be associated with attrition. This allows us to concentrate on semester-to-semester associations between an individual's STEM-grading deficit and persistence in STEM among the (policy-relevant) group of *all* first-year STEM-actives. Furthermore, we expand on existing relative grading research by estimating the role of a "STEM grading penalty" as a push factor (away from STEM), net of absolute grading (in college), and academic preparation (before college).

5 | RESULTS

5.1 | Exodus and influx into STEM

Consistent with the argument by Cannady et al. (2014) that multiple routes towards a STEM degree can be found outside of the modal STEM pipeline, we first examine the consequences of switching college majors. To what extent do initial STEM students leave a STEM program?

Table 1 reports the enrollment size by intended STEM major in 2004 (column 1) and the percentage of students within each major who later graduated with a BA in Health (column 2) *or* with a non-STEM degree (column 3) by 2009, while excluding non-graduates. Overall, these percentages indicate a rather small amount of

TABLE 1 STEM exodus to non-STEM fields between first enrollment in 2004 and 2009: College graduates only

	First-semester enrollments (2004)	Percent health BA in 2009	Percent other non-STEM BA in 2009
Intended STEM major as freshman (2004)			
Biological/biomedical sciences	250	4.8%	35.7%
Computer/information science	120	0.0%	49.2%
Engineering	290	1.4%	22.0%
Mathematics/statistics	40	0.0%	29.5%
Physical sciences	60	4.8%	25.8%
	760	2.5%	31.5%
Taking STEM in 2004 (no STEM major declared)	2,260	10.6%	70.3%
Total students active in STEM	3,020		

Note: $N = 3{,}020$ (all graduates active in STEM in semester 1). The columns titled "percent health BA in 2009" and "percent other non-STEM BA in 2009" are mutually exclusive. Its sum represents the total share of non-STEM majors among graduates in the Beginning Postsecondary Students Longitudinal Study.

leaking out of STEM due to competition from health-related majors (2.5% on average), as well as a considerable amount of leaking into other non-STEM majors: 31.5% on average. Together, this means that a total of 260 students in the sample left a declared STEM major before graduation.

If we look at students who had *not* declared a college major as a freshman but nonetheless took STEM coursework during the first semester, the share who became non-STEM graduates is larger: about 80% (bottom row, Table 1). We define this nontraditional group of first-year STEM students as either taking one or more nonremedial STEM courses or taking at least four nonremedial STEM credits. Although a much smaller percentage of those non-declaring majors graduate college with a STEM-degree, about 430 (rounded) are among this sizeable group of students (19% of 2,260) who took STEM from the start of their college careers.

In addition to outflows, we look for a possible *STEM-influx* from non-STEM programs over the course of a student's college years. Table 2 indicates, for each intended non-STEM major in the first semester, the number of students enrolled (column 1), the percentage of students who had graduated with a degree in STEM in 2009 (column 2), and the percentage of students who were also taking STEM coursework in their first semester (column 3).

We observe that a considerable proportion of students who initially enroll in technical or semitechnical programs eventually graduate with a STEM degree. This is far above the average of 10.6% STEM influx across all majors. Interestingly, about 18% of first-year health-related students eventually transition into a STEM major and graduate within 6 years, indicating that the suggested leakage of STEM into health programs is not nearly as widespread as the reverse trajectory. Furthermore, students who have expressed no interest in a specific major in the first semester (labeled "undeclared/undecided") or are initially enrolled in liberal arts ultimately produce a relatively high percentage of STEM graduates: 15% and 10%, respectively.

By comparing this influx of 490 students (a rounded figure) to the earlier STEM exodus (260 students) in the nationally representative BPS sample, we conclude that the number of students *moving into STEM* during college is considerably larger than the number of students leaving STEM. These numbers call into doubt the imagery of a "leaky STEM pipeline," because the reverse effect—students moving into STEM from other disciplines—predominates.

5.2 | STEM-actives

On the basis of the widespread involvement with STEM coursework across the freshman population and the substantial influx later on into STEM from other college majors, we believe it is necessary to reconsider the

TABLE 2 Science, technology, engineering, and mathematics (STEM) influx from non-STEM fields between first enrollment in 2004 and 2009

Intended major in 2004	Enrollments (2004)	Percent STEM BA (2009)	Percent taking STEM in first semester[a]
Technology-related	60	83.1%	92.3%
Construction trades	10	44.4%	33.3%
Science technology	10	30.8%	92.3%
Health professions	460	17.9%	79.0%
Undeclared/undecided	1,670	15.4%	47.8%
Liberal arts/Humanities	90	10.1%	38.2%
Agriculture	50	9.3%	72.2%
Architecture	40	7.5%	32.5%
Transportation	30	7.4%	81.5%
Psychology	170	5.8%	39.9%
Visual/performing arts	230	4.4%	21.2%
Education	450	3.8%	37.2%
Languages/Linguistics	30	3.6%	28.6%
Parks/Leisure/Fitness studies	30	3.6%	53.6%
Business, management, Marketing	640	2.8%	44.5%
Philosophy/Theology	50	2.1%	31.9%
Homeland security	100	2.1%	20.8%
Communication/Journalism	230	1.8%	32.5%
Social sciences	230	1.7%	33.2%
Other non-STEM majors[b]	140	0.0%	68.1%
Total/mean	4,720	10.6%	45.7%

Note: $N = 8{,}980$ (all first-time freshmen), 5,500 graduated. Students who did not graduate within 6 years after enrollment were excluded.
[a]Took at least four nonremedial STEM credits or took at least one nonremedial STEM course.
[b]A group of small majors with no graduates in STEM.

definition of a "first-year STEM student." A large group of potential STEM graduates is simply not sure about their academic goals at college entry. Their ambition for STEM fields may be latent or may develop later on in college, depending on their experiences, performances, and other developing interests.

From this point on, we will concentrate on a more-inclusive subsample of first-year students that combines those students who indicated an interest in STEM (surveyed in the first semester) with those who took at least one nonremedial STEM course (or four credits) in the first semester ($N = 3{,}022$). Some non-STEM programs require a nonremedial course in math or science early on, which means that "taking at least one STEM course" is not strictly—or not always—an active decision by the student. One potential limitation of this approach is that, with regard to STEM-actives outside of a STEM program, we cannot be sure whether this signifies an early interest or simply checking off a requirement. As known from the literature, early STEM-taking could indicate an existing interest in STEM, as well as *create* an interest in STEM among students outside of the traditional STEM majors (Maltese et al., 2014). We, therefore, argue that

TABLE 3 Science, technology, engineering, and mathematics (STEM) retention by type of first-semester "STEM-actives"

	Percent STEM BA in 2009
First-semester STEM major	
Biological and biomedical sciences	59.5%
Computer/information science	50.8%
Engineering	76.6%
Mathematics/statistics	70.5%
Physical sciences	69.4%
Total average	**66.0%**
First-semester non-STEM major (but STEM-active)	
Health	17.6%
Other	15.5%
Undecided	24.8%
Total average	**19.1%**

Note: N = 3,020 students (all STEM-actives, non-graduates excluded).

STEM-actives should be considered potential college STEM graduates by researchers and policymakers who are concerned with STEM dropout.

Almost 30% of all STEM-active students in our sample eventually graduated with a BA degree in STEM. However, as documented in Table 3, the breakdown by choice of major in the first semester shows substantial variability in STEM graduation rates. The top-panel indicates the percent of students who graduated in STEM (out of all students who graduated) by the first-semester STEM field. Similarly, the bottom panel lists the STEM graduation percentages for STEM-actives who enrolled in a health major, a different non-STEM major, or who were undecided in their first semester. The retention among STEM-actives (bottom panel) is much higher for those students who declared a traditional STEM major at the beginning of college compared to students who did not, but were nonetheless involved with STEM coursework. Among students who declare a STEM major in the first college semester 66% of them eventually graduate with a BA in STEM.

Although the variation in attrition among different STEM majors is not the main aim of this paper, computer and information science majors were noteworthy for having somewhat lower retention rates (50%) compared to the overall STEM major average. It should be noted that first-semester computer science majors have a slightly lower level of STEM involvement early on vis-à-vis the other declared STEM majors; seven versus nine STEM credits, respectively. Furthermore, compared to the other traditional STEM majors, the average GPA in both STEM and non-STEM coursework is slightly lower for computer science majors, as well as the gap between STEM grades and non-STEM grades in favor of the latter. These are relevant components of attrition mechanisms among the population of interest, "STEM-actives," as discussed later.

In contrast, students who declare an interest in STEM, those who do not indicate a STEM intention during their first semester—but who are nevertheless active in STEM coursework—display lower overall STEM graduation rates. As shown in the bottom panel of Table 3, about 19% of students in health, other majors, or who remain undecided, graduate with a STEM degree. Although the graduation rate among these STEM-actives is lower compared to early declared STEM majors, it is remarkable that a quarter of students who did not indicate a specific degree preference *at all* (i.e., the first-semester "major undecided" group) eventually graduate with a BA degree in STEM. This percentage is higher compared to students who attempt a health-related major, which is known for its partial overlap with STEM coursework.

Taken together, these descriptive data show that STEM graduates come from a much wider range of first-semester college students. We also conclude that students with a plan to major in STEM in their first semester—those typically included in STEM pipeline research—remain much more likely to persist in STEM programs. We will,

therefore, control for "major intention" (STEM or non-STEM) as a control variable in our regression analyses, as well as the number of STEM credits attempted per semester.

5.3 | Grading

One important factor in determining whether STEM-actives continue in STEM is their academic performance, specifically their grades in STEM compared to in their other courses. Figure 1 reports the average semester grade for both STEM graduates and non-STEM graduates in the eight semesters after their first enrollment, split by students' STEM and non-STEM GPA. Across all groups, students seem to adjust to coursework; on average, students receive higher grades in their eighth semester than at the beginning of their college career. STEM graduates generally receive *higher* grades compared to their non-STEM graduating counterparts. This suggests that STEM graduates possess strong academic skills that benefit their coursework in both STEM and non-STEM courses.

Figure 1 also reveals that STEM courses are harder in terms of their grading than non-STEM courses, for all groups. Indicated by the two solid lines, both STEM graduates (rectangular markers) and non-STEM graduates (triangle markers) display a STEM GPA that is roughly 3–4 decimal points lower than their own non-STEM GPA. Although average grades of STEM courses fluctuate slightly and display a small improvement between the first semester and the eighth semester, the gap between those observations and the non-STEM averages remains remarkably stable across consecutive semesters. We interpret this as STEM being graded tougher *relative* to non-STEM courses at any stage of the college career, even among the academically strongest (STEM)-graduating students.

At the same time, comparing the two dotted lines also indicates that students who graduated in STEM (rectangular markers) scored higher in non-STEM coursework than their non-STEM graduating counterparts (triangle markers) throughout the college career. In other words, those who ultimately graduate in STEM indicate an academically stronger record than non-STEM graduates, regardless of the type of coursework. This suggests that predictive models of STEM graduation should control for ability measures and prior academic performance to correctly estimate the net effects of grading or other covariates.

Figure 2 plots the "STEM-penalty" across the first eight semesters among STEM-actives. This STEM-penalty measures the difference between a student's non-STEM GPA and that same student's STEM GPA (subtracting the former from the latter, each semester). This difference is the "GPA points difference" or 'STEM penalty marked on

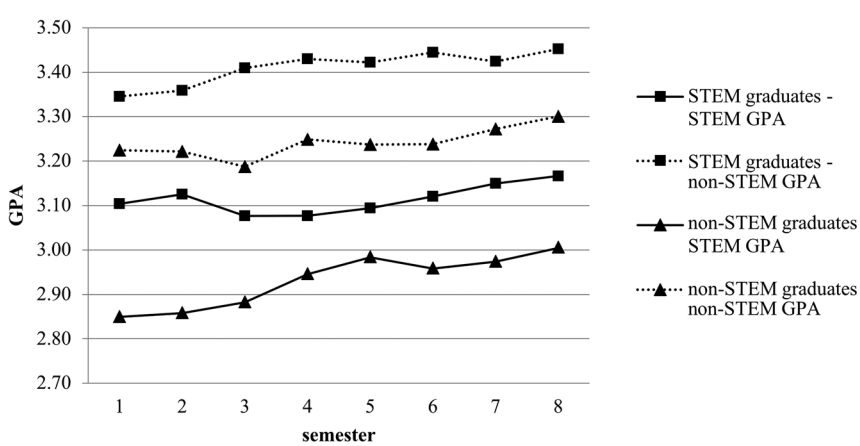

FIGURE 1 Grading in STEM across eight semesters among STEM-actives (N = 3,020; all graduates). GPA, grade point average; STEM, Science, technology, engineering, and mathematics

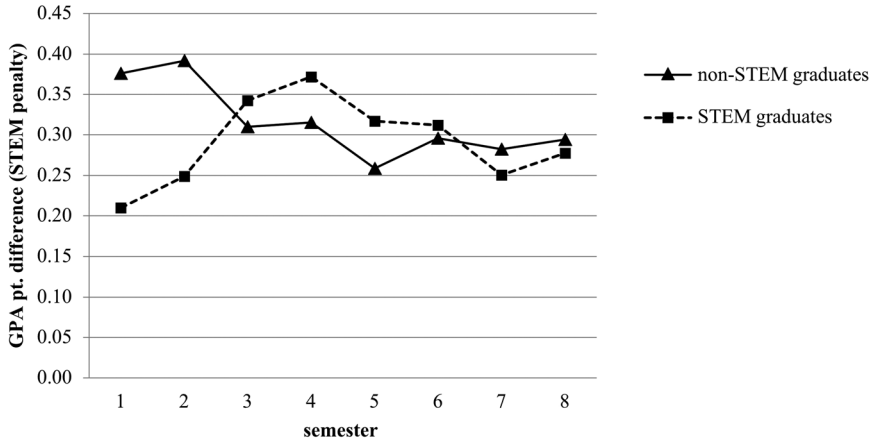

FIGURE 2 Development of the Science, technology, engineering, and mathematics (STEM) grading penalty among STEM-actives by semester ($N = 3{,}020$; all graduates). T-statistics STEM versus non-STEM differences: Semester 1 (−9.89), semester 2 (−7.82); both $p < .001$. Among STEM graduates, the uptick between semester 1 versus semester 4 is statistically significant ($t = 270.81$, $p < .001$)

the *y*-axis of Figure 2. As seen in this figure, all STEM-active students suffer grade-wise from taking STEM; the average difference between one's STEM GPA and one's non-STEM GPA is consistently in favor of the latter. By taking STEM coursework, students pay a price in terms of their average grades, which are roughly 0.25 to almost 0.4 GPA decimal points lower in STEM courses. As the STEM-penalty only refers to relative grading, this gap may in theory also be caused by students who score disproportionally high in non-STEM coursework or disproportionally low in STEM coursework. However, the relative grading trajectories of Figure 2 provide a clear indication that regardless of persisting and graduating in STEM, taking STEM courses means opting for a more difficult path towards graduation.

Furthermore, Figure 2 shows that the STEM penalty follows a remarkable pattern for STEM- and non-STEM graduates. STEM-actives who graduate with a non-STEM major display a significantly higher STEM-penalty in the first few semesters of college compared to their counterparts who obtained a STEM degree. First-year STEM students who later switch to a non-STEM major faced a bigger gap between their STEM and non-STEM course grades. Thereafter, their penalty slowly decreases as these students presumably take fewer or easier STEM courses.

The pattern is very different, however, for those who do graduate with a STEM degree, indicated by the striped line. They initially receive a significantly ($p < .001$) lower STEM-penalty in the first couple of semesters compared to non-STEM graduates, indicating that they perform relatively well in STEM courses. Yet in subsequent college semesters STEM graduates suffer from a higher and consistently increasing grade penalty. More specifically, their average STEM-penalty increases significantly between semester 1 and semester 4, from 0.207 to 0.371 GPA points. The average grade penalty remains high throughout the college career but declines slightly in the last semesters which may suggest an adjustment to STEM coursework.

Given prior research demonstrating that women persist in STEM at much lower rates compared to men, we also calculate the STEM-penalty by gender. These indicate that, on average, women receive a statistically significantly higher STEM-penalty than men in the first couple of semesters: 0.354 versus 0.230 in semester 1 and 0.313 versus 0.212 in semester 2, respectively (not shown). The multivariable analyses of STEM graduation presented later will help clarify whether relative grading also operates differently for men and women in relation to persisting in STEM.

Taken together, Figure 2 points at the potential role of relative grading in explaining and predicting the STEM leakage. Low early grades in STEM relative to one's other courses may push some STEM-actives away from STEM

(and into other majors). For first semester STEM-active students who persist in taking STEM courses, however, pursuing a STEM degree means going through a "STEM boot camp": a trajectory that starts with accepting a considerable GPA penalty in one's STEM courses, followed by a steep increase of this penalty for another 2 years of their college career.

5.4 | Predicting STEM graduation

How important is this grading penalty for explaining why first-year STEM-active students leave STEM? We estimate this relationship using logistic regression models that are clustered on the first four college semesters (Table 4), nested into three models according to their theoretical impact. All cases selected as "STEM-actives" are included in each of these models. In addition to the logit coefficient, we report the average marginal effect (AME) for each significant predictor; the change in the probability of STEM graduation when the independent variable increases by one unit, conditional on all other predictors. The focal predictor of STEM-graduation in our models—the STEM penalty, measured as an individual's non-STEM GPA subtracted from his/her STEM GPA—is calculated for the first four semesters separately.

Model 1 includes demographic background factors as well as the educational experience predictors that have previously been associated with academic success in STEM. Our model confirms that men are, on average, 15.9%-points more likely to graduate with a STEM degree than women. Asians have a higher chance of STEM graduation (AME of +10.3%-points), whereas students whose native language is English graduate less frequently. As found in earlier studies, high school performance is a strong predictor of success for STEM graduates: a high GPA (3.5+) and taking (pre-)calculus benefit STEM-active students.

Model 2 adds the first-semester intention to major in a STEM field, with the "undeclared majors" as the reference group. This model confirms that declaring STEM is positively associated with graduation in a STEM field. After controlling for demographics and precollege predictors, first-year STEM-declared students are 35.6%-points more likely to graduate in STEM, in comparison to the "undecided" group. STEM-active premedical students are slightly less likely to graduate in a STEM field compared to the undecided reference group.

Model 3 adds several proximal causes of STEM attrition, whereby the negative association of withdrawing from a STEM course (−5.2 percentage points) and failing a STEM course (−3.4 percentage points) are expected. However, most striking are the added semester-by-semester STEM penalties; indicating that, after controlling for other factors (including STEM GPA), receiving a STEM-penalty in the first or second semester is associated with greater attrition. These marginal effects are relatively small, −2.4 percentage points and −2.0 percentage points, but are statistically significant. In other words, regardless of educational background, academic performance, and STEM "ambition," and one's STEM performance, receiving lower STEM grades relative to one's non-STEM grades at the beginning of one's college career is associated with the exit of STEM-active students out of STEM BA programs.

However, as visible in Figure 2, the grading penalties of STEM graduates steadily increase between semester 1 and 4 (a "STEM boot camp"), whereas those who leave STEM receive a much lower penalty after the first couple of semesters. Presumably, many STEM-active students leave STEM *before* semester 4, which explains why a STEM-penalty in semester 4 is statistically associated with graduation (a margin of +3.7 percentage points). We argue that this parabolic shaped effect of semester-by-semester relative grading not only indicates the significant role that grades play in explaining why ambitious STEM freshmen leave these programs quickly but also shows that those who continue to take STEM coursework are consistently disadvantaged in their GPA.

5.5 | Heterogeneity and robustness checks

In this study, we cannot observe all components of motivation and academic ability that may be associated with lower STEM "ability" or STEM graduation chance. However, we control for a range of measures that should capture

TABLE 4 Clustered (4-semesters) logistic regression on STEM graduation among STEM-actives

	Model 1		Model 2		Model 3	
	Logit	Margin	Logit	Margin	Logit	Margin
Gender						
Male	0.859***	15.9%	0.550***	8.3%	0.530***	6.7%
Race[a]						
Black	0.312~	5.9%	0.109		0.284	
Latino	0.110		−0.022		0.034	
Asian	0.526**	10.3%	0.439*	7.0%	0.230	
Other	0.237		0.242		0.178	
Age	0.000		0.019		0.021	
English native speaker	−0.559***	−10.3%	−0.528**	−8.0%	−0.581***	−7.4%
High school GPA[b]						
Low (0–1.9)	0.628		0.595		1.079	
Medium-low (2–2.9)	−0.467**	−8.6%	−0.537*	−8.1%	−0.452*	−5.8%
High (3.5–4)	0.409***	7.6%	0.364**	5.5%	0.308*	3.9%
High school math[c]						
None	0.607*	10.5%	0.679*	10.1%	0.663*	8.4%
Algebra I/II	−0.167		−0.058		0.086	
Pre-calculus	0.375**	6.2%	0.428**	6.1%	0.424**	5.2%
Calculus	1.110***	21.2%	0.964***	14.9%	0.896***	11.6%
Intended major first semester[d]						
STEM			1.687***	35.6%	1.361***	23.2%
Pre-med			−0.331*	−5.2%	−0.653***	−9.0%
Other			−0.721***	−10.3%	−0.760***	−10.3%
English remedial					−0.568***	−7.2%
Math remedial					0.101	
Withdrew STEM					−0.411***	−5.2%
Failed STEM					−0.265*	−3.4%
STEM credits attempted[e]						
0–2.99					−2.201***	−28.0%
3–5.99					−0.853***	−13.5%
>10					0.571***	10.3%
STEM GPA[f]						
A					0.018	
B					−0.234*	−3.0%
D					−0.747***	−9.0%
STEM penalty (GPA)[g]						
Semester 1					−0.192**	−2.4%
Semester 2					−0.157*	−2.0%
Semester 3					0.130	1.6%
Semester 4					0.289***	3.7%
Summary statistics						
N	3,020		3,020		3,020	
R^2	0.100		0.231		0.342	

Note: *N* = 3,020 students (all STEM-actives). Variables added in Model 3 are measured across the first four college semesters (clustered standard errors). All variance inflation factors remain close to 1 and not higher than 1.3, indicating no concern for multicollinearity. Reference: [a]White; [b]medium-high (3–3.4); [c]trigonometry; [d]undecided; [e]6–9.99 credits; [f]C, range; [g]GPA difference between non-STEM coursework and STEM coursework calculated in such a way that a positive value indicates a "penalty."

*$p < .05$.

**$p < .01$.

***$p < .001$.

these mechanisms, such as high school GPA, high school math preparation, remedial coursework taken, and absolute STEM grades in college—all as indicators of ability (Table 4). Our main model also controls for a key motivation indicator: intended major in the first college semester. Nonetheless, after adjusting for these factors, a STEM penalty (relative grading) still proves to significantly predict attrition from STEM activity. We, therefore, consider this relative grading effect to reflect a psychological discouragement rather than a true underperformance.

Sample sizes allow us to check whether the negative effect of a STEM-penalty on STEM graduation holds for particular subgroups of STEM-actives, such as those who did not declare a STEM major in the first semester (column 1) and those two only took two or more STEM courses in the first semester (column 2). If the narrowing STEM pipeline is just a consequence of "weeding out"—a selection of only the highest achievers in math or the most STEM-committed students—we would not see a significant impact of a STEM-penalty among these STEM-active students. However, as shown in the third column of Table 5, we still find evidence for a significant impact of a STEM-penalty on STEM graduation even among the *most* committed STEM students—those who had indicated their STEM intentions at the beginning of college. When we disaggregate the analyses even further by running the models for each first-semester-declared STEM major, we find that the models suffer from lack of statistical power. They are, therefore, not shown. Nonetheless, the negative STEM-penalty point estimates on STEM persistence are observed in all five traditional STEM majors subsamples. We, therefore, believe that the relative grading mechanism operates in a similar fashion among various early declared STEM majors.

In addition, based on the higher average level of the STEM grading penalty among women compared to men, we run a series of predictive models of STEM graduation with interactions between each of the semester's STEM penalties and gender (not shown). These models indicate nonsignificant interactions between gender and the STEM-penalty variable. The implication is that even though women receive a higher STEM-penalty than men, neither gender is disproportionally affected by it with regard to the chance to graduate in a STEM field.

Finally, we test a scenario whereby the significant impact of the STEM-penalty on STEM attrition is driven by students receiving a STEM-penalty as a result of being a high achiever in non-STEM coursework (while also being a strong student in STEM courses). However, as shown in the rightmost column of Table 5, even if we exclude the top-20% performing students in non-STEM coursework across the first couple of semesters, the estimates of receiving a STEM-penalty in these semesters remain statistically significant and negatively associated with STEM graduation.

TABLE 5 STEM-penalty estimates on alternative subsamples

	STEM-taking, but not STEM-declared		Two or more STEM courses in semester 1		STEM-declared only		Excluding top-20% performers in non-STEM	
	Logit	Margin	Logit	Margin	Logit	Margin	Logit	Margin
STEM penalty (GPA)								
Semester 1	−0.284**	−3.1%	−0.214*	−3.3%	0.024	0.4%	−0.225**	−2.8%
Semester 2	−0.037	−0.4%	−0.336**	−5.1%	−0.362**	−6.3%	−0.172*	−2.2%
Semester 3	0.223*	2.4%	0.046	0.7%	0.029	0.5%	0.087	1.1%
Semester 4	0.314**	3.4%	0.195	3.0%	0.204	3.6%	0.288**	3.6%
Summary statistics								
N	2,210		1,320		780		2,410	
R^2	0.246		0.313		0.176		0.324	

Note: Subsamples (rounded) of all STEM-actives 3,020 students. Variables are measured across the first four college semesters (clustered standard errors). Estimates are adjusted for all predictors from Model 3 in Table 4.
*$p < .05$.
**$p < .01$.
***$p < .001$.

In other words, though a STEM-penalty may reflect being a strong student in the arts or humanities, and therefore, being pulled into these fields, later on, this mechanism does not influence the identified relationship between relative grading and STEM attrition.

Aside from these robustness checks regarding ability and motivation, the literature has identified several potentially observable factors associated with STEM attrition that unfortunately cannot be accounted for in the current study due to data availability. These include the type of institution, level or type of college counseling, the demography of course instructors (gender and race), and details regarding high school AP courses.

6 | DISCUSSION

At the entry to 4-year college education, there is a sizeable group of students who are involved in STEM. This group is much larger than just the students who clearly state their STEM ambition in the first semester—the latter group for which the "leaky pipeline" is often studied and understood by policymakers (Metcalf, 2010). Building on a research framework that has emphasized the fact students may declare a STEM interest late or switch majors based on new preferences after college entry (Maltese et al., 2014; Nauta, 2007), we identify a substantial group of "STEM-actives" who attempt STEM credits and courses in their first semester. We argue that using this broader benchmark for evaluations of STEM attrition processes is informative for both educational researchers and policymakers who seek to increase the number of STEM graduates. Non-STEM-enrolled students who take non-remedial STEM coursework in their first semester either do so voluntarily or as a program requirement. Our definitional choice cannot distinguish the precise context of students' STEM involvement and thereby the strength of students' interest in STEM at this early stage in college. However, STEM-active students who are not enrolled in a STEM program may be trying out STEM or—in case of a requirement—may become interested in STEM. We find a wide range of students for whom a STEM degree might be realistically pursued, either through a strictly defined STEM track or through the accumulation of coursework and switching from other majors.

We argue that *in addition to* the first semester declared STEM students, STEM-active students should be viewed as the wider group among which potential STEM graduates should be sought, and among which the phenomenon of leaving STEM needs to be studied. Although the group traditionally analyzed (STEM-declared students) may consist of more motivated and better-prepared students, we have shown that the other STEM-active students display surprisingly high STEM graduation rates. To be precise, our analysis indicates that a substantial number of nondeclared yet STEM-taking students *still* graduate in a STEM field within 6 years after initial college enrollment. Moreover, we find that an influx into STEM exceeds the outflow from STEM into other majors, calling into question the very notion of STEM leakage *in college*. The relevance of studying all STEM-actives rather than STEM-declared students only is also implicitly supported by research based on large-scale survey data. Maltese et al. (2014) showed that STEM graduates often report an early interest in STEM, but that the triggering of STEM interest can happen throughout the college career. Furthermore, the fact that a substantial proportion of STEM graduates are merely STEM course-taking, rather than declared STEM students, in their early college semesters, confirms Cannady et al.'s (2014) multilayered STEM pipeline.

Regardless of major influx and exodus, educational researchers and policymakers may still ask why some college students who indicate an interest in STEM or who are involved in substantive STEM coursework nevertheless graduate with a different non-STEM major. In this study, we concentrate on the association between grading patterns in the early stages of the college career and subsequent STEM attrition in the following semesters. We draw on an existing framework in educational research that has shown students' future course selection is partly a function of push and pull signals drawn from their relative performance in prior coursework (Bar et al., 2009; Fournier & Sass, 2000).

When concentrating on the first-year STEM-active students, we observe the expected effects of student background, high school performance and preparation, and college-related factors (STEM GPA and STEM credits taken)

in our regression models, as known from prior research. Subsequently, we observe an *additional* process that is associated with leaving STEM: a relationship between experiencing a STEM grade penalty in the first couple of semesters and graduating with a BA degree in STEM. The significant relevance of a STEM grade penalty for STEM attrition—the difference between each student's STEM GPA and his/her non-STEM GPA within a particular semester—is observable even among otherwise highly talented four-year college students. After controlling for precollege GPA, STEM GPA intended major, and mathematical preparation, lower relative grades in STEM courses in the first couple of semesters are *still* negatively associated with STEM retention. This leads us to conclude that a STEM-penalty—the impact of relative grading over and above absolute grading and preparation—has a negative influence on STEM graduation chances.

These results, drawn from a representative sample of all US college freshman, are in line with some previous results from a single liberal arts college that had first identified an association between relative grading in specific courses (departments) and persistence in these STEM fields (Rask & Tiefenthaler, 2008; see Rask, 2010 for caveats). Our findings expand on this knowledge by showing both the existence of a STEM-wide grading penalty among all STEM-actives and its development across several semesters. Importantly, the predictive models indicate a significant negative association between receiving this penalty within the first couple of semesters and attrition from a STEM field (i.e., graduating in non-STEM field).

More generally, these findings demonstrate that relative grades are an important proximal cause of strong students dropping out of STEM in college. These STEM attrition signals should not be regarded as simply weeding out the low-performing students among STEM-actives. A thorough assessment of the negative impact of the STEM-penalty on STEM graduation rules out that lower academic performance in STEM or STEM prerequisites forms the underlying mechanism of attrition. The most elaborate predictive model adjusts for many academic factors including high school math, declaring STEM early, math remediation requirements, and college STEM GPAs.

Our analysis also suggests that there is heterogeneity in who receives a STEM-penalty. For instance, even though the identified mechanism of relative grading on STEM graduation does not operate differently across gender, women suffer from higher STEM penalties than men early on in their college careers. We, therefore, believe that future research should concentrate on the forms of social heterogeneity in relative grading that currently remain beyond the scope of this study: the discrepancy in exposure among different groups of students and the institutional contexts within which STEM penalties occur. Furthermore, receiving substantially higher grades in non-STEM coursework may cause a STEM-penalty to arise even if a student scores above average in STEM coursework. Robustness checks show that this scenario does not influence the relative grading effect on STEM attrition. It remains plausible that higher performance in non-STEM courses operates as a pull factor (into non-STEM majors) rather than a push factor (away from STEM) among strong STEM-actives. Further disentangling these mechanisms will require different—qualitative—data on students' incentives.

One limitation of the proposed relative grading mechanism for a leaking STEM pipeline is that, almost by definition, it remains a partial explanation; many other structural factors still matter. However, based on this study, we can draw two specific questions and a directional hypothesis for follow-up research and policy design. Our analyses reveal that STEM courses are graded lower level *relative* to non-STEM coursework and that this gap is significantly associated with students moving away from STEM coursework early on in the college career. This means that if one seeks to improve STEM persistence, one essential evaluation should first concentrate on the reasons behind grading processes by STEM instructors and STEM departments more generally. To what extent are these related to the content of instruction or to the comparison to non-STEM teaching? If such evaluations allow for the STEM grading distribution to be shifted or expanded—allowing higher grades for the best performances—one hypothesis based on the current study would suggest a positive effect on the STEM graduation rate. Subsequently, of importance in the policy evaluations of such new practices or experiments is the question of which groups of students experience a shift in their STEM penalty (a possible narrowing of the STEM grades and the non-STEM grades). Thus, aside from the fact that we know some students experience a higher

STEM penalty (e.g., women), one needs to know which groups would benefit from new grading practices to understand the total effect on STEM persistence.

ACKNOWLEDGMENTS

We thank the anonymous reviewers of Science Education for their insightful comments and suggestions. These helped improve the manuscript substantially.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

ORCID

Dirk Witteveen http://orcid.org/0000-0002-9197-8316

REFERENCES

Alper, J., & Gibbons, A. (1993). The pipeline is leaking women all the way along. *Science*, *260*, 409–411.

Altonji, J. G., Blom, E., & Meghir, C. (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annual Review of Economics*, *4*, 185–223. https://doi.org/10.1146/annurev-economics-080511-110908

Aughinbaugh, A. (2012). The effect of high school math curriculum on college attendance: Evidence from the NLSY97. *Economics of Education Review*, *31*(6), 861–870. https://doi.org/10.1016/j.econedurev.2012.06.004

Bar, T., Kadiyali, V., & Zussman, A. (2009). Grade information and grade inflation: The Cornell experiment. *Journal of Economic Perspectives*, *23*(3), 93–108. https://doi.org/10.1257/jep.23.3.93

Becker, H. S., Geer, B., & Hughes, E. C. (1968). *Making the grade: The academic side of college life*. New York, NY: John Wiley & Sons Inc.

Bettinger, E. (2010). To be or not to be: Major choices in budding scientists. In C. T. Clothfelter (Ed.), *American universities in a global market* (pp. 69–100). Chicago, IL: University of Chicago Press.

Blickenstaff, J. C. (2005). Women and science careers: Leaky pipeline or gender filter? *Gender and Education*, *17*(4), 369–386. https://doi.org/10.1080/09540250500145072

Bottia, M. C., Stearns, E., Mickelson, R. A., Moller, S., & Valentino, L. (2015). Growing the roots of STEM majors: Female math and science high school faculty and the participation of students in STEM. *Economics of Education Review*, *45*, 14–27. https://doi.org/10.1016/j.econedurev.2015.01.002

Cannady, M. A., Greenwald, E., & Harris, K. N. (2014). Problematizing the STEM pipeline metaphor: Is the STEM pipeline metaphor serving our students and the STEM workforce? *Science Education*, *98*(3), 443–460. https://doi.org/10.1002/sce.21108

Carnegie Corporation of New York, & Institute for Advanced Study. (2009). *The opportunity equation. Transforming mathematics and science education for citizenship and the global economy*. New York, NY: Carnegie Corporation of New York.

Carnevalle, A.P., Smith, N., & Melton, M. (2011). *STEM. Georgetown University Center on Education and the Workforce* (Report ED525297). Retrieved from https://files.eric.ed.gov/fulltext/ED525307.pdf

Catsambis, S. (1994). The path to math: Gender and racial-ethnic differences in mathematics participation from middle school to high school. *Sociology of Education*, *67*(3), 199–215. https://doi.org/10.2307/2112791

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin*, *135*(2), 218–261. https://doi.org/10.1037/a0014412

Chen, X., & Soldner, M. (2013). *STEM attrition: College students' paths into and out of STEM fields. National Center for Education Statistics (2014–001)*. Washington, DC: National Center for Education Statistics.

Chizmar, J. F. (2000). A discrete-time hazard analysis of the role of gender in persistence in the economics major. *The Journal of Economic Education*, *31*(2), 107–118. https://doi.org/10.1080/00220480009596768

Correll, S. J. (2001). Gender and the career choices process: The role of biased self-assessments. *American Journal of Sociology*, *106*(6), 1691–1730. https://doi.org/10.1086/321299

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

Daempfle, P. (2002). *An analysis of the high attrition rates among first year college science, math and engineering majors*, Washington, DC: U.S. Dept. of Education, Office of Educational Research and Improvement, Educational Resources Information Center (ERIC).

Diekman, A. B., Brown, E. R., Johnston, A. M., & Clark, E. K. (2010). Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science*, *21*(8), 1051–1057. https://doi.org/10.1177/0956797610377342

Fairweather, J. (2009). *Linking evidence and promising practices in science, technology, engineering and mathematics (STEM) undergraduate education* (Status Report). National Academies National Research Council Board of Science Education. Retrieved from https://www.nsf.gov/attachments/117803/public/Xc–Linking_Evidence–Fairweather.pdf

Federman, M. (2007). State graduation requirements, high school course taking, and choosing a technical college major. *The B.E. Journal of Economic Analysis & Policy*, *7*(1), 32. https://doi.org/10.2202/1935-1682.1521

Fischer, S. (2017). The downside of good peers: How classroom composition differentially affects men's and women's STEM persistence. *Labour Economics*, *46*(June), 211–226. https://doi.org/10.1016/j.labeco.2017.02.003

Fournier, G. M., & Sass, T. R. (2000). Take my course, *please*: The effect of the principles experience on student curriculum choice. *The Journal of Economic Education*, *31*(4), 323–339. https://doi.org/10.1080/00220480009596450

Gayles, J. G., & Ampaw, F. (2014). The impact of college experiences on degree completion in STEM fields at four-year institutions: Does gender matter? *The Journal of Higher Education*, *85*(4), 439–468. https://doi.org/10.1080/00221546.2014.11777336

Gehrke, S., & Kezar, A. (2017). The role of STEM faculty communities of practice in institutional and departmental reform in higher education. *American Educational Research Journal*, *54*(5), 803–833. https://doi.org/10.3102/0002831217706736

Geiger, R. (1997). What happened after Sputnik? Shaping university research in the United States. *Minerva*, *35*(4), 349–367. https://doi.org/10.1023/a:1004324331995

Goulden, M., Frasch, K., & Mason, M. A. (2009). *Staying competitive: Patching America's leaky pipeline in the sciences*. Oakland, CA: Center for American Progress, Berkeley Center on Health, Economic, & Family Security and the Center for American Progress, University of California. http://cdn.americanprogress.org/wp-content/uploads/issues/2009/11/pdf/women_and_sciences.pdf

Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review*, *29*(6), 911–922. https://doi.org/10.1016/j.econedurev.2010.06.010

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology*, *94*(3), 562–575. https://doi.org/10.1037//0022-0663.94.3.562

Hill, C., Corbett, C., & St. Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics*. Washington, DC: American Association of University Women.

Hrabowski, F. A., III. (2012). Broadening participation in the American STEM workforce. *BioScience*, *62*(4), 325–326. https://doi.org/10.1525/bio.2012.62.4.2

Hurtado, S., Han, J. C., Sàenz, V. B., Espinoza, L. L., Cabrera, N. L., & Cerna, O. S. (2007). Predicting transition and adjustment to college: Biomedical and behavioral science aspirants' and minority students' first year of college. *Research in Higher Education*, *48*(7), 841–887. https://doi.org/10.1007/s11162-007-9051-x

Jaggars, S. S., & Hodara, M. (2011). *The opposing forces that shape developmental education: Assessment, placement, and progression at CUNY community colleges* (CCRC Working no. 36). New York, NY: Columbia University: Teacher's College.

Kokkelenberg, E. C., & Sinha, E. (2010). Who succeeds in STEM studies? An analysis of Binghamton University undergraduate students. *Economics of Education Review*, *29*(6), 935–946. https://doi.org/10.1016/j.econedurev.2010.06.016

Long, M. C., Conger, D., & Iatarola, P. (2012). Effects of high school course-taking on secondary and postsecondary success. *American Educational Research Journal*, *49*(2), 285–322. https://doi.org/10.3102/0002831211431952

Main, J. B., Mumford, K. J., & Ohland, M. W. (2015). Examining the influence of engineering students' course grades on major choice and major switching behavior. *International Journal of Engineering Education*, *31*(6), 1468–1475.

Malcom, L. E., & Dowd, A. C. (2008). The impact of undergraduate debt on the graduate school enrollment of Latino STEM baccalaureates. *The Review of Higher Education*, *35*(2), 265–305. https://doi.org/10.1353/rhe.2012.0007

Maltese, A. V., Melki, C. S., & Wiebke, H. L. (2014). The nature of experiences responsible for the generation and maintenance of interest in STEM. *Science Education*, *98*(6), 937–962. https://doi.org/10.1002/sce.21132

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students. *Science Education*, *95*(5), 877–907. https://doi.org/10.1002/sce.20441

Marra, R. M., Rodgers, K. A., Shen, D., & Bogue, B. (2012). Leaving engineering: A multi-year single institution study. *Journal of Engineering Education*, *101*(1), 6–27. https://doi.org/10.1002/j.2168-9830.2012.tb00039.x

Metcalf, H. (2010). Stuck in the pipeline: A critical review of STEM workforce literature. *InterActions: UCLA Journal of Education and Information Studies*, *6*(2). Retrieved from https://escholarship.org/uc/item/6zf09176

Miller, D. I., & Wai, J. (2015). The bachelor's to Ph.D. STEM pipeline no longer leaks more women than men: A 30-year analysis. *Frontiers in Psychology*, *37*(6), 1–10. https://doi.org/10.3389/fpsyg.2015.00037

Morgan, S. L., Gelbgiser, D., & Weeden, K. A. (2013). Feeding the pipeline: Gender, occupational plans, and college major selection. *Social Science Research*, *42*(4), 989–1005. https://doi.org/10.1016/j.ssresearch.2013.03.008

National Academy of Sciences, National Academy of Engineering, & Institute of Medicine. (2007). *Rising above the gathering storm*. Washington, DC: The National Academies Press. https://doi.org/10.17226/11463

National Center for Education Statistics. (2011). *2004/2009 Beginning Postsecondary Students Longitudinal Study Restricted Use Data File*, Washington, DC: US Department of Education.

National Science Board. (2012). *Science and engineering indicators 2012*. Arlington, VA: National Science Foundation (NSB 12-01).

National Survey of Student Engagement (NSSE). (2013). *A fresh Look at student engagement* (Annual Report). Indiana University Center for Postsecondary Research. Retrieved from http://nsse.indiana.edu/NSSE_2017_Results/pdf/NSSE_2017_Annual_Results.pdf

Nauta, M. M. (2007). Assessing college students' satisfaction with their academic majors. *Journal of Career Assessment*, *15*(4), 446–462. https://doi.org/10.1177/1069072707305762

Nicholls, G. M., Wolfe, H., Besterfield-Sacre, M., Shuman, L. J., & Larpkiattaworn, S. (2007). A method for identifying variables for predicting STEM enrollment. *Journal of Engineering Education*, *96*(1), 33–44. https://doi.org/10.1002/j.2168-9830.2007.tb00913.x

Ohland, M. W., Sheppard, S. D., Lichtenstein, G., Eriz, O., Chachra, D., & Layton, R. A. (2008). Persistence, engagement, and migration in engineering programs. *Journal of Engineering Education*, *97*(3), 259–278. https://doi.org/10.1002/j.2168-9830.2008.tb00978.x

Ong, M., Wright, C., Espinosa, L., & Orfield, G. (2011). Inside the double bind: A synthesis of empirical research on undergraduate and graduate women of color in science, technology, engineering, and mathematics. *Harvard Educational Review*, *81*(2), 172–208. https://doi.org/10.17763/haer.81.2.t022245n7x4752v2

Ost, B. (2010). The role of peers and grades in determining major persistence in the sciences. *Economics of Education Review*, *29*(6), 923–934. https://doi.org/10.1016/j.econedurev.2010.06.011

Peri, G., Shih, K., & Sparber, C. (2015). STEM workers, H-1B visas, and productivity in US cities. *Journal of Labor Economics*, *33*(2), 225–255. https://doi.org/10.1086/679061

President's Council of Advisors on Science and Technology: PCAST (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Executive Office of the President of the United States. Retrieved from https://files.eric.ed.gov/fulltext/ED541511.pdf

Price, J. (2010). The effect of instructor race and gender on student persistence in STEM fields. *Economics of Education Review*, *29*(6), 901–910. https://doi.org/10.1016/j.econedurev.2010.07.009

Rask, K. (2010). Attrition in STEM fields at a liberal arts college: The importance of grades and pre-collegiate preferences. *Economics of Education Review*, *29*(6), 892–900. https://doi.org/10.1016/j.econedurev.2010.06.013

Rask, K., & Tiefenthaler, J. (2008). The role of grade sensitivity in explaining the gender imbalance in undergraduate economics. *Economics of Education Review*, *27*(6), 676–687. https://doi.org/10.1016/j.econedurev.2007.09.010

Rask, K. N., & Bailey, E. M. (2002). Are faculty role models? Evidence from major choice in an undergraduate institution. *The Journal of Economic Education*, *33*(2), 99–124. https://doi.org/10.1080/00220480209596461

Redmond-Sanogo, A., Angle, J., & Davis, E. (2016). Kinks in the STEM Pipeline: Tracking STEM graduation rates using science and mathematics performance. *School Science and Mathematics*, *116*(7), 378–388. https://doi.org/10.1111/ssm.12195

Riegle-Crumb, C., & King, B. (2010). Questioning a white male advantage in STEM: Examining disparities in college major by gender and race/ethnicity. *Educational Researcher*, *39*(9), 656–664. https://doi.org/10.3102/0013189X10391657

Riegle-Crumb, C., King, B., Grodsky, E., & Muller, C. (2012). The more things change, the more they stay the same? Prior achievement fails to explain gender inequality in entry into STEM college majors over time. *American Educational Research Journal*, *49*(6), 1048–1073. https://doi.org/10.3102/0002831211435229

Robinson, M. (2003). Student enrollment in high school AP sciences and calculus: How does it correlate with STEM careers? *Bulletin of Science, Technology & Society*, *23*(4), 265–273. https://doi.org/10.1177/0270467603256090

Sabot, R., & Wakeman-Linn, J. (1991). Grade inflation and course choice. *Journal of Economic Perspectives*, *5*(1), 159–170. https://doi.org/10.1257/jep.5.1.159

Sadler, P. M., Sonnert, G., Hazari, Z., & Tai, R. (2014). The role of advanced high school coursework in increasing STEM career interest. *Science Educator*, *23*(1), 1–13.

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave sciences*. Boulder, CO: Westview Press.

Shaw, E. J., & Barbuti, S. (2010). Patterns of persistence in intended college major with a focus on STEM majors. *The National Academic Advising Association Journal*, *30*(2), 19–34. https://doi.org/10.12930/0271-9517-30.2.19

Simpson, J. C. (2001). Segregated by subject: Racial differences in the factors influencing academic major between European Americans, Asian Americans, and African, Hispanic, and Native Americans. *Journal of Higher Education*, *72*(1), 63–100. https://doi.org/10.1080/00221546.2001.11778865

Stinebrickner, R., & Stinebrickner, T. R. (2014). A major in science? Initial beliefs and final outcomes for college major and dropout. *The Review of Economic Studies*, *81*(1), 426–472. https://doi.org/10.1093/restud/rdt025

Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. A. (2011). STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology*, *100*(2), 255–270. https://doi.org/10.1037/a0021385

Tai, R. H., Liu, C. Q., Maltese, A. V., & Fan, X. (2006). Planning early for careers in science. *Life Science*, *312*(5777), 1143–1144.

Tyson, W. (2011). Modeling engineering degree attainment using high school and college physics and calculus coursetaking and achievement. *Journal of Engineering Education*, *100*(4), 760–777. https://doi.org/10.1002/j.2168-9830.2011.tb00035.x

Tyson, W., Lee, R., Borman, K. M., & Hanson, M. A. (2007). Science, technology, engineering, and mathematics (STEM) pathways: High school science and math coursework and postsecondary degree attainment. *Journal of Education for Students Placed at Risk*, *12*(3), 243–270. https://doi.org/10.1080/10824660701601266

Wang, M., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields. *Developmental Review*, *33*(4), 304–340. https://doi.org/10.1016/j.dr.2013.08.001

Whalen, D. F., & Shelley, M. C., II. (2010). Academic success for STEM and non-STEM majors. *Journal of STEM Education: Innovations and Research*, *11*(1), 45–60.

Wine, J., Janson, N., & Wheeless, S. (2011). *Beginning postsecondary students longitudinal study (BPS:04/09)* (Methodology Report no. NCE 2012-246). Washington, DC: National Center for Education Statistics.

Xie, Y., Fang, M., & Shauman, K. (2015). STEM Education. *Annual Review of Sociology*, *41*, 331–357. https://doi.org/10.1146/annurev-soc-071312-145659

Xie, Y., & Killewald, A. (2012). *Is American science in decline?* Cambridge, MA: Harvard University Press.

Xue, Y., & Larson, R. C. (2015). STEM crisis or STEM surplus? Yes and yes. *Monthly Labor Review*, *U.S. Bureau of Labor Statistics*, 1–14. https://doi.org/10.21916/mlr.2015.14

How to cite this article: Witteveen D, Attewell P. The STEM grading penalty: An alternative to the "leaky pipeline" hypothesis. *Science Education*. 2020;104:714–735. https://doi.org/10.1002/sce.21580