Passing Through the Gates:
Identifying and Developing Talent in Introductory STEM Courses

M. Kevin Eagan*, Gina A. Garcia, Sylvia Hurtado, Josephine A. Gasiewski

University of California, Los Angeles

*Contact: M. Kevin Eagan, 405 Hilgard Ave., 3005 Moore Hall, University of California,

Los Angeles, CA 90095-1521; Phone: (310) 825-1925; Email: mkeagan@gmail.com

This study was made possible by the support of the National Institute of General Medical Sciences, NIH Grant Numbers 1 R01 GMO71968-01 and R01 GMO71968-05, the National Science Foundation, NSF Grant Number 0757076, and the American Recovery and Reinvestment Act of 2009 through the National Institute of General Medical Sciences, NIH Grant 1RC1GM090776-01. This independent research and the views expressed here do not indicate endorsement by the sponsors.

Introduction

Introductory science, technology, engineering, and mathematics (STEM) courses often function to eliminate all but the "top tier" students and champion the idea that "scientists are born, not made" (Tobias, 1990, p.11). Critics argue that introductory courses encourage passive learning (i.e., Bransford et al., 2000), with the majority of attrition from the sciences occurring within the first two years of taking these courses (Chang et al., 2008). Few studies, however, have looked at the positive outcomes associated with these courses. In this mixed methods study, we examine alternative measures of talent and determine how it is developed and harvested within introductory STEM courses.

Students with aspirations for STEM bachelor's degrees encounter significant obstacles in the form of "gatekeeper" courses almost as soon as they begin their collegiate coursework. Success in introductory science and math coursework represents the necessary first step toward the completion of a bachelor's degree in STEM, as introductory courses provide foundational learning for all further coursework. Unfortunately, rather than providing tools for future study, many introductory courses tend to discourage many students from continuing in the sciences (Seymour & Hewitt, 1997).

Instructors typically base grades in introductory science courses on students' ability to acquire and retain specific content knowledge rather than on their development of critical thinking skills, the latter of which are equally as necessary as the former for future science careers (Gainen, 1995). By assessing students primarily on mastery of content knowledge, gatekeeper courses resemble sorting mechanisms that eliminate potentially talented students before their talent can be fully realized and developed. Several scholars have concluded that prior academic achievement largely determines students' grades in introductory coursework

(Gainen, 1995; Kamii, 1990; NCES, 2000; Payzant & Wolf, 1993; Waits & Demana, 1988); therefore, gatekeeper course grades may be more of a reflection of students' prior abilities than the acquisition of knowledge and skills during the course. Particularly in the case of students of color who have traditionally be underrepresented in STEM (as well as women), prior achievement, motivation, and socioeconomic status represent key predictors of success in college science and engineering courses (NCES, 2000; Payzant & Wolf, 1993; Waits & Demana, 1988).

Success in introductory courses is theoretically a function of both scientific thinking dispositions and content knowledge (Conley, 2005; Hagedorn, Siadat, Fogel, Nora, & Pascarella, 1999). Unfortunately, in practice, college-level introductory science and mathematics courses tend to focus too much on the acquisition of content knowledge and too little on the development of meta-cognitive skills related to critical thinking and scientific literacy (Handelsman, Ebert-May, Belchner, Bruns, Chang, DeHaan, Gentile, Lauffer, Stewart, Tilghman, & Wood, 2004; Hurd, 1997; Williams, Papierno, Makel, & Ceci, 2004). As a result, introductory science courses tend to have relatively high failure and dropout rates (Seymour & Hewitt, 1997). Failure to succeed in gatekeeper courses can lead to difficulty in future courses and may prompt students to switch out of science majors (Labov, 2004; Seymour, 2001).

Building on previous research this study takes a broad view of success that encompasses not only the narrow construct of students' course grade but also other conceptualizations of talent and performance, including thinking and acting like a scientist. We hypothesize that student success will be affected by a variety of factors that can broadly be broken down into four conceptual areas: students' prior preparation in sciences; students' experience with the learning environment in their introductory courses; instructors' pedagogy; and the amount of effort students expend on the course.

We conceive of the student experience with the course learning environment as being multifaceted, encompassing how students feel about and interpret their instructors' teaching methods, coursework, and learning environment. We conceptualize course pedagogy as the kind of instructional methods that are used, the emphasis of the coursework, and the type of learning environment constructed by the instructor. Student effort, background characteristics, and initial propensity toward thinking and acting like a scientist are included in the model because these factors have all been shown to affect student success in coursework; controlling for these variables allows for the examination of the unique impact of introductory courses and other related experiences on student success.

Methodology

In order to evaluate alternative measures of talent in introductory STEM courses, we utilized a sequential explanatory mixed methods design, which consisted of collecting, analyzing, and integrating both quantitative and qualitative data during the research process (Creswell, 2005; Teddlie & Tashakkori, 2009). For a full description of our mixed methods approach, please see Gasiewski, Eagan, Garcia, Hurtado, and Chang (2012). The data utilized in this study are part of a larger, multi-phased project sponsored by the National Institutes of Health (NIH) focusing on innovative techniques in introductory STEM courses.

Quantitative Analyses

Data and Sample. During the spring of 2010, we surveyed students and faculty in introductory STEM courses at 15 institutions across the U.S. The sample of institutions was relatively diverse and included three historically Black colleges and universities (HBCUs), three Hispanic-serving institutions (HSIs), eight public institutions, one technical university, and two liberal arts institutions. Within each institution, we sampled an average of five introductory

STEM courses. We defined introductory as the first course in a sequence of courses where knowledge is cumulative. Each institution provided at least one biology course and at least one chemistry course. We also had a mix of introductory calculus, statistics, engineering, and computer science courses across all of our participating institutions; however, the distribution of the types of courses represented varied across each institution.

At the beginning of the academic term, students in these courses completed the web-based 2010 STEM Student Pre-Questionnaire, which collected information on students' self-rated academic and science abilities, the frequency with which they articulate and apply science concepts, and demographic information. Many of the items on the pre-survey were based on Conley's (2005) work in identifying the skills and dispositions students should have for success in introductory science courses. During the last two weeks of the academic term, students completed the web-based 2010 STEM Student Post-Questionnaire, which re-asked many of the same questions from the first survey while also including a number of items related to students' experiences in their introductory courses. Faculty who taught these courses completed an online instructor survey at the end of the course, which included items related to the pedagogical techniques faculty used in the course, their perceptions of student learning, and their priorities for undergraduate education. Finally, we collected course grade information from registrar's offices at each of our institutions. In all, we had 3,205 students across 76 classrooms in 15 institutions respond to both student surveys, which translated into a 42.1% response rate. Gasiewski et al. (2012) provides additional information about the survey administration and the specific items included on each survey.

Measures. This study focuses primarily on three outcome variables, measured at or near the end of the academic term. Two outcomes are latent constructs that represent the frequency with which students reported acting and thinking like scientists, and these same two constructs were measured as pre-tests in the first survey in order to control for prior abilities and experiences. The third outcome measure is a variable representing students' final grade in their introductory course. The two latent constructs, frequency of thinking like a scientist and acting like a scientist, were composed of a set of indicator variables identified through factor analysis. A confirmatory factor analytic model was run in MPlus to confirm that the relationships between the indicator variables and the constructs held up for both the pre-test and the post-test. Table 1 presents the indicator variables explained by the latent constructs, their factor loadings, and the fit indices for the measurement model. The appendix presents a full list of all of the predictor variables included in the analyses, including demographic characteristics, pre-college preparation, and experiences in introductory STEM courses.

Analyses. Through multilevel structural equation modeling (SEM), we analyzed relationships among exogenous and endogenous variables in an effort to simultaneously estimate the relationships among sets of variables and confirm latent constructs (Bentler, 2006; Bentler & Wu, 2002). Parameter estimates are generated by analyses of estimated covariance matrices. SEM accounts for measurement error and provides overall goodness of fit indices to determine the adequacy of the model, both of which represent advantages over traditional path analysis (Laird, Engberg, & Hurtado, 2005). Using a multilevel structural model accounts for the nested nature of the data, as students are clustered within classrooms. This technique also helps us to avoid making a Type I statistical error. To assess model fit, we relied on two fit indices: comparative fit index (CFI) and root mean square error of approximation (RMSEA). CFI values above 0.90 indicate adequate model fit, while RMSEA scores below 0.06 indicate an appropriate level of fit (Raykov, Tomer, & Nesselroade, 1991).

Our analytic approach began with a confirmatory factor analysis that tested the adequacy of our measurement model. As mentioned above, the measurement model included the observed indicator variables and their associated latent constructs for both the pre- and post-surveys. This measurement model confirmed the factor structure of the two pre-test factors and their associated post-tests. Next, we added to the measurement model all of the hypothesized predictors and paths to test the full structural model. LaGrange Multiplier tests provided guidance, in conjunction with prior literature and theory, about adding relational paths among variables in the model. If all paths from a variable were removed, we dropped the variable from the analysis.

Qualitative Analyses

Data and Sample. From the 15 institutions that participated in the quantitative survey, we purposefully selected eight institutions based upon high levels of classroom innovation occurring on their campus and conducted interviews and focus groups with faculty and students. The eight universities were located across the United States and included: one Hispanic serving institution (HSI), one historically Black college/university (HBCU), and six predominantly White institutions (PWI). Of these institutions, four are publically funded. Gasiewski et al. (2012) provides more information about the characteristics of these institutions.

In sum, 41 focus groups were conducted over a five-month time span, from October 2010 to February 2011, with 239 student participants from the eight universities. The student focus groups consisted of students enrolled in introductory STEM courses between Spring 2010 and Spring 2011. Interviews ranged from 60 to 90 minutes and were conducted with two to ten participants per session, averaging five focus groups per campus. We asked students to describe their experiences in introductory STEM courses through a series of nine main questions and corresponding probes, centering around student motivation, course structure, learning, instruction, and assessment, allowing their responses to dictate the order with which we asked the questions.

In addition to the student focus groups, we conducted 25 in-depth interviews with faculty members across the eight institutions. Interviews were conducted during the same time span as the focus groups. Every faculty member associated with the introductory courses sampled for the study was invited to participate, yet only 25 individuals agreed to be interviewed. Thus our faculty sample is based upon self-selection. Faculty members in the sample teach introductory courses in various disciplines including chemistry (n=10), biology (n=9), mathematics (n=5), and engineering (n=1). Years of teaching ranged from two for the newest faculty members to forty years for the most experienced. Through a series of seven main questions and corresponding probes, we asked them to describe their introductory STEM course including goals and objectives, pedagogical approaches, structure, forms of assessment, and institutional support for teaching.

For both in-depth interviews and student focus groups, we utilized a semi-structured interview technique that allowed us to respond "to the situation at hand, to the emerging worldview of the respondent, and to new ideas on the topic" (Merriam, 1998). Maxwell (2005) suggests that this technique increases the "internal validity and contextual understanding and is particularly useful in revealing the processes that led to specific outcomes" (p. 80). All interviews were digitally recorded, transcribed verbatim by a professional transcription company, checked for accuracy, and loaded into NVivo8 qualitative software.

Analyses. In order to develop the coding architecture utilized in NVivo, each transcript was open coded by examining the raw data and identifying salient themes supported by the text. This constant comparative approach followed an inductive process of narrowing from particular (text segments) to larger themes while allowing the researcher to attempt "to 'saturate' the categories—to look for instances that represent the category and to continue looking until the new information does not provide further insight into the category" (Creswell, 2007, pp. 150-151). Our team of six researchers each read transcripts from two institutions, gathering and comparing themes across focus groups/interviews and institutions, which also enabled analytical triangulation (Patton, 2002). Once we determined that we had reached saturation in generating themes, we developed several iterations of coding schemes, wherein codes were created, expanded, defined, and refined. After bins of relevant data were created, we re-read the data repeatedly in order to solidify our understanding and see connections amongst the categories.

Limitations

A number of constraints with the methodology and the data may limit the generalizability of our findings and conclusions. For the quantitative portion of our study, we situated our study within 76 classrooms across 15 campuses of varying type, size, selectivity, and mission, and the composition of our sample may limit the generalizability of our findings to other types of institutions and classroom contexts. Additionally, because this study examined students' experiences in a single introductory course, the short timeframe in which we administered the surveys may have affected the amount of change detected in students' frequency of thinking and acting like a scientist. This study chose to focus on student changes in these items over the span of a single academic term; had students been tracked over a longer period of time, we may have detected more substantial changes in these outcomes.

For the qualitative data collection, it is critical to note that the students who participated in the focus group interviews were not necessarily enrolled in the classes taught by the professors who agreed to be interviewed. Because of this limitation, we cannot make direct comparisons between faculty and students but rather begin to uncover the teaching and learning context at each institution more broadly based on both groups' experiences. Finally, we realize that interviewing only 3-4 faculty members per institution limits our ability to discuss faculty perspectives; however, we used the data to begin to uncover some of the elements necessary for understanding introductory STEM classrooms within different contexts.

Results

Alternative Ways to Identify Talent

The first step in our analysis was to determine if, beyond grades, there are sufficient ways to determine talent in introductory STEM courses. Using student-level quantitative data, we specifically looked at latent constructs that represent students' ability to think like a scientist and act like a scientist. Table 1 presents the results of the measurement model for the four latent variables in the study: pre- and post-tests of thinking like a scientist and acting like a scientist. The model statistics suggest that the model fits the data well, as the CFI was 0.96 and the RMSEA was 0.03. Table 1 shows the factor loadings of observed variables on each of the four factors; all variables loaded highly on the relevant factor. The post-test latent constructs represent alternative measures of talent. Acting like a scientist includes items related to the frequency with which students conducted experiments, synthesized information from multiple sources, and identified scientific articles and resources. Thinking like a scientist includes indicators of making predictions based on existing knowledge, asking relevant questions in class, making sense of scientific concepts, and translating scientific language into non-scientific terminology.

In addition to these latent constructs, we specifically asked faculty members in our study to tell us how they identify talent in introductory STEM courses. Two main themes emerged from the data that are similar to the two identified constructs, specifically the one related to thinking like a scientist. Although faculty were not in total agreement about the importance of grades in determining talent, they agreed that asking and answering good questions reveals that a student can make sense of the concepts, think critically about the material, and translate the material into language that is understandable to other students (i.e., thinking like a scientist).

Grades *do not* matter. For some faculty members teaching introductory STEM courses, grades are a better indicator of how well a student can memorize information than how talented he/she is. Additionally, performance may be an indicator of hard work, time on task, and preparation. One faculty member said that some students have figured out the best way to navigate tests and others have learned how to maximize partial credit given on exams. In this case, grades may not necessarily mean that a student is particularly gifted in the subject. Professor Lewis from Northeastern Private Master's College argued that, since his quizzes draw straight from the homework, performance is not an indicator of talent.

> The quizzes I ask in calculus are typically drawn straight from the homework problems. They're two homework problems verbatim just repeated where I got them. All it shows is preparation, [so] that's a pretty good mark of —if they can understand something enough to kind of spit it back out at you in a shorter timeframe.

On the contrary, students who do not perform well on exams may actually be very talented. In this case, faculty members take note of other qualities that the student possesses, as suggested by Professor Pace at Western Public Research University:

> Yeah. I had a student…he got [a] B plus, but he would solve problems that nobody could solve. He wouldn't be able to solve the problems that everybody could solve, but he solved the problems that no one could. Now, that was very impressive, but he didn't do well on the exams…he actually did very well later on.

As suggested by this quote, talented students often stand out in other ways, despite performance on tests. Two professors, in particular, talked about persistence, maturity, and motivation as important traits that talented students possess.

> I keep track of their performance, not only on exams. I also follow the students through the semester in terms of how they have been improving. Maybe, let's say, I start with somebody that made a 50 on my first exam, but then that person ends up with an 80 something on the last exam, so you see progress. Okay. You see that the person has not given up. You see that the person is actually working hard. That persistency sometimes is also a key, so I look at that persistence. (Professor Shepard, Southwestern Public Research University)

> I had [a student] for two semesters and over the course of the year I saw him adapt, mature, do successively better on all his exams. I saw that he was interested in—that he seemed serious about a career in the sciences and so I was—when he expressed an interest in doing [a summer research program] I was really enthusiastic about recommending him. The feedback I got from the woman who runs the program after it was over was that he had done really well. There was a case where he appeared to struggle in the beginning but I was able to see over time that he was actually a pretty committed and hard working young man and I saw a lot of maturity just over the course of the year. (Professor Hume, Southeastern Public Master's College)

Instead of determining a student's talent based on performance on one or two exams, these faculty members have chosen to watch how students develop over time. Another way that faculty said they gain a better sense of students' talent is to take time to get to know them in smaller settings (i.e., lab). By talking to students, professors said they can determine whether a student is a good test taker or truly interested in the subject. The smaller setting also allows faculty the opportunity to have more focused conversations with students that go beyond the class lecture.

Grades *do* matter. On the contrary, some faculty members teaching introductory STEM courses believe that performance on tests is the best way to identify talent. Professor Burke at Southwestern Private Research University, for example, stated the following:

> Certainly by their outstanding performance on my exams. If I give a hard exam and they do really, really well, then I know they're really sharp. At one point I used to have access to—the fact that they're in the honors class means they have a superior SAT or ACT test score.

A faculty member at Western Public Research University had a similar answer, implying that the talented students will inevitably stand out based solely on performance while another at

Midwestern Public Research University said there is absolutely no way to know who is talented prior to the first exam.

> The only way for me to—there's no way for me to—okay. In any given lecture, by the time I'm two or three weeks in, there will be five or six students who answer the majority of the questions. I'll start to recognize those students. If a student is quiet and does not come to office hours, there's absolutely no way I would know how good that student is doing until the first exam, so that's the truth of it. (Professor Dawson, Midwestern Public Research University)

For some faculty members, performance is the best way to identify a student for a research opportunity or for an opportunity as a Supplemental Instructor facilitator. If a student does exceptionally well throughout the term, some faculty may approach the student about changing his/her major or about considering a minor. A few faculty members went a step further and explained that by doing well on an exam, a student can demonstrate that he/she has the ability to apply the concepts, critically analyze it, and synthesize it. These skills, which are similar to the latent construct of thinking like a scientist, may ultimately translate it into a research environment, where students can begin to act like scientists.

> They have to have a good grasp of not just the content but the ability to apply it. I kind of have, I think, a reputation as being very rigorous and expecting students to go beyond just rote memorization, and so I look for students who have that ability because in many cases it does translate into the research environment. Being able to ask questions, being able to take information and critically analyze it, figure out how you're gonna alter an experiment. You know, how to interpret your data. All of that really comes from utilization of information, not just memorize it, stick it up there, and hope it stays. (Professor Smith, Western Private Master's College)

Asking and answering questions. For faculty who look beyond performance and grades, another way to identify talent is to listen to the students who ask questions in class. Not only do talented students ask questions but they ask thoughtful questions that indicate that they are processing the concepts being presented. One professor said that talented students do more than just memorize the step-by-step procedures while another indicated that talented students do more than just solve the problems. These skills become evident in the types of questions students ask.

> For me, okay, in math, I'm not very impressed by the ability of working problems out 'cause you can memorize that, but if suddenly you see a student asking a question that [makes me wonder], "Where is that coming from?" So it's asking questions. In lower division class if suddenly they will ask you a question that you would just completely be, "How did he or she think of that?" (Professor Pace, Western Public Research University)

The type of questions students ask, therefore, is an important way for undergraduate students in STEM to distinguish themselves in an introductory course by showing that they can think like scientists. Students who connect the concepts in a STEM course with those in a social science course and those who question the information being presented show signs of talent.

> By those people who ask good questions in class. By people who are not willing to accept things on face value. Or they will dispute statements that I make. Sometimes they win a few, sometimes they lose a few, sometimes we get a few rained out. I like that. (Professor Burke, Southwestern Private Research University)

As suggested by Professor Burke, the talented students not only stand out in class but also invoke excitement within the professor. Professor Veerdansky at Western Private Master's College, for example, became excited as he told the following story about a student who he identified as "really good."

> I like the questions they ask, so for the vert bio I'll be lecturing long and I'll ask a little question here and there that might be pointed. You know, like, "how do you think the sharks ventilate if they're not doing this buccal pumping kind of thing, cuz they don't have the operculum?" I'll get them to, I answer questions in class just to make sure you're kinda tracking me or thinking about stuff. But then the ones that I'm like, whew, you're really good, are, "Okay, you've told me about how they change their osmoregulation when they go from fresh water to salt water. How exactly does that happen, and how does it happen on the way back?"

Faculty also suggested that the way students answer a question is an important indicator of talent. Answering questions in class tells the professor how deeply the student is thinking and whether he/she is beginning to think like a scientist. For some professors, this may be more important than test results because some students are naturally good at taking tests and memorizing the material. Professor Locke (Western Public Research University) said that "people pleasers" will shout out an answer without thinking through the process while those with science talent will take a moment to answer and then provide support for their answer:

> Every time I've recognized somebody that I think has really good research capabilities, thought processes or—what do they call them competency in that way of thinking, they'll always give their idea and then tell me why.

Beyond providing support for their answers, talented students in science also find the answers themselves by looking up scientific articles (one indicator of acting like a scientist), as suggested by Dr. Verdansky (Western Private Master's College).

> This kid in particular…she goes back and she looks up [a] scientific paper. Sends me the scientific paper saying, "Can you look at – this is what I found. This is the answer to my question that I found." That, to me, says there is some major intellectual curiosity. So thinking about a good question and then actually going to the place that you need to figure it out, and then sharing that back and really continuing that discussion. But you know, it stands out.

Although faculty suggested that one way to identify talent is by noticing the questions that students are asking and answering in class, it may be difficult for some students to speak up in class. Professor Thomas and Professor Reyes suggested that another way to identify talented students is to watch their interaction with other students.

> The other thing I do though is I watch for—you can see even in the big groups, you can see the student walk in and all of a sudden three students ask him questions. You watch for those students, the ones that kinda seem to attract others, because they're giving the answers and they may not be the A, B student, but they're giving the answers and they're working towards it. Somehow they have a personality of leader type and so you watch those students as well, so you just kinda watch the dynamics. (Professor Reyes, Southwestern Public Research University)

This quote suggests that not only are faculty able to distinguish the students with talent but other students are also likely to notice when a peer is exceptional. Professor Reyes said that students may also stand out within the Blackboard online course website as those that answer other students' questions and organize study groups. Together, both student-level quantitative data and faculty level qualitative data revealed that, beyond grades, there are important indicators of talent within the STEM disciplines including acting and thinking like a scientist and asking/answering questions in a way that displays knowledge and understanding at a deeper level. We tested this idea further by looking at predictors of grades in introductory STEM courses as well as predictors of acting and thinking like a scientist.

Factors Predicting Grades and Other Measures of Talent

The second step in our analysis was to determine the relationship between the latent constructs and the final grades that students received. More specifically, we were interested in whether alternative measures of talent could be used to predict final grades in introductory STEM courses. Table 2 shows the results of the final structural model, including unstandardized regression coefficients, standardized regression coefficients, standard errors, and significance levels for the model's direct effects. Figure 1 diagrams the causal paths in the final structural model. The fit indices for the final model included a CFI of 0.94 and an RMSEA of 0.04, which demonstrated adequate fit.

Final grade. Our analyses suggest that students' final grades do not relate to alternative measures of student talent, as the results in Table 2 suggest that thinking and acting like a scientist were not significantly correlated with students' final grades in their introductory STEM courses. Instead, two of the most important predictors of students' grades were students' composite SAT scores and their grades in high school biology. These findings connect to prior research that demonstrates that students with higher SAT scores and greater levels of academic achievement in high school typically earn significantly higher grades early in college compared to their peers with lower test scores and high school grades (Astin, 1993; Sharkness et al., 2011).

Similarly, students who reported feeling well-prepared for this course tended to earn significantly higher grades than their counterparts who did not feel well prepared.

Among the variables corresponding to students' experiences in the course, findings in Table 2 suggest that students who felt more confident in their ability to learn when they began the introductory course tended to earn significantly higher grades than students who lacked this confidence. Additionally, students who considered dropping the course earned significantly lower end-of-course grades than their peers who never considered dropping the course. The standardized coefficients suggest that considering dropping the course represented the most important predictor of students' final course grade. This finding connects to students' confidence in their learning ability, as students who may have picked up on early cues that their performance, as measured by grades, was lacking. Likewise, changing study habits during the academic term corresponded to significantly lower final grades, as students likely changed how they prepared for the course due to low scores on early exams and assignments. Having higher self-rated time management skills predicted significantly higher end-of-course grades, which again connects academic performance to college ready skills.

Data from the student focus groups suggest that having confidence in introductory STEM courses is related to high school preparation. For this reason, it is not surprising that SAT scores, high school biology grades, confidence, and higher self-rated time management skills predicted the final grade. Several students mentioned that the AP courses they took in high school were more challenging than their introductory STEM courses.

> But like, high school, my high school, I went to a high school up north. And this has – these courses are like, as rigorous as my high school AP courses. (Aidan, Southwestern Private Research University)

Other students said they thought their college courses would be difficult because their high school teachers told them that college courses would be more challenging. For these students, introductory STEM courses seemed easier than expected, thus reinforcing their confidence. Rosie suggested that the tests in high school were more challenging because they included short answers questions, which require students to think more critically about the material than multiple choice tests.

> Personally, I thought they'd be a little more difficult I want to say. My biology class, it was a lot different than the one I took in high school and it was all multiple choice and there were application questions but in high school we had like short answers where you had to write out problems and draw like a chromosome and draw this and label that. I felt more challenged in that class than this one but I still learned a lot. (Rosie, Western Private Master's College)

Both quantitative and qualitative student-level data, therefore, reinforce the idea that high school preparation and performance positively predict student outcomes in introductory STEM courses.

In looking at how faculty pedagogical strategies in introductory courses predict students' end-of-course grades, the path from the faculty variable of encouraging collaboration among students to grades was not significant. One predictor, taking a course where faculty members more frequently relied upon essay exam questions, did significantly predict lower course grades. This finding may signal that students struggle when having to demonstrate extended knowledge of a concept compared to what would be expected from a multiple choice exam. With essay questions, students must apply their knowledge and demonstrate critical thought to arrive at the "correct" answer. A quote from Miranda supports this conclusion.

> Then, so then when it came to the test you would study from the book but, and from your notes, but what was on the test was really – like what he would do is teach you so much and then he wanted you to apply that on a test more than test you on what you knew. He wanted you to apply it, which was a little harder 'cuz I wasn't an expert, so for me that was a little difficult. (Miranda, Southeastern Private Master's College)

Multiple students at all eight campuses admitted to memorizing for the test as opposed to learning the underlying concepts of the course. The tendency of faculty to use multiple choice tests in introductory STEM courses enables students to rely on rote memorization, which may ultimately hinder their success. When students were asked what they might do differently if given the chance to take introductory STEM courses again, Becca said the following.

> I think – I guess trying to learn more, understanding the concept versus learning for the test. Not so much like knowledge that I'm kind of expected to like, "Well, you took discrete math, you should know how to do all this," and it's like, when I took the class, you know, I got an A in the class and I learned what I was going to be tested on. But then after the class, it is kind of – I don't like, I didn't care about it as much so I didn't remember all the details, and then I'm coming to other classes, that they expect that I – it's like a pre-req, so they expect that you know all this stuff. You kind of come into problems where you don't remember it. (Becca, Northeastern Private Master's College)

This quote further supports what some faculty said about grades not mattering when it comes to identifying talent. Just because Becca got an A in her introductory math course does not mean that she retains information in math. In thinking about ways to foster the development of talent, faculty need to consider ways to improve students' conceptual understanding and retention of course material.

Thinking like a scientist. Looking beyond grades, we further examined predictors of our alternative measures of talent. The findings in Table 2 provide details on the characteristics and behaviors that contribute to students' increased frequency of thinking and acting like scientists. As expected, students' frequency of thinking like a scientist at the beginning of the academic term served as the strongest predictor of this latent construct at the end of the term. The second most important predictor was students' confidence in their ability to learn, as students who felt more confident in their learning ability tended to report that they more frequently thought like a scientist. Unlike the results for final course grade, neither SAT scores nor high school biology grade had a direct relationship with students' frequency of thinking like a scientist. The results of the model suggest a significant difference in thinking and acting like a scientist between White and non-White students, as White students reported thinking like a scientist significantly more often than their peers at the end of the course.

Among the course-related predictors, four variables had a significant relationship with students' frequency of thinking like a scientist. Students who more frequently attended review sessions also reported thinking like a scientist significantly more often than their peers who attended review sessions less often. This is similar to earlier findings by Gasiewski et al. (2012) that revealed that students who attended supplemental instruction sessions were more engaged in the course than students who did not attend. Being in a course that emphasized applying concepts to new situations significantly and positively predicted students' frequency of thinking like a scientist, which may speak to students engaging with course material in new and interesting ways. Likewise, students who felt academically prepared for their introductory STEM courses reported significantly higher scores on the construct of thinking like a scientist compared to their peers who felt less prepared. Students who considered dropping their introductory STEM course during the academic term had significantly lower scores on thinking like a scientist.

The importance of feeling academically prepared was also prevalent in the qualitative data, with students indicating that doing well in high school biology and chemistry helped them do well in college-level courses. Doing well in introductory STEM courses, however, may also help students feel prepared for the next level while enhancing their ability to think like scientists.

> Well, I took Basic Chemistry last year, and I'm taking General Chemistry, which is the next step above it, and I feel like I was really prepared for it. 'Cuz right now I'm in Gen Chem [and] like, I already know this, yeah? Like, I guess the professor who taught me was good at what she was doing 'cuz I already knew what I was doing and like, right now some kids are already confused about like, the stuff we learned last year. And we were supposed to know this already, but I guess they were confused because of the professor. But for me it was kind of a breeze. (Sameer, Southwestern Public Research University)

Only one classroom-level variable had a significant relationship with thinking like a scientist. Students who were in classrooms where faculty felt that there was no question that was too elementary reported significantly higher scores on thinking like a scientist. Again, this finding connects to earlier work by Gasiewski et al. (2012) that found students become more engaged in introductory STEM courses when faculty signal an openness to questions. Students seem to pick up on these accessibility cues and become more willing to engage in activities related to thinking like a scientist.

Acting like a scientist. Linked to thinking like a scientist was the frequency with which students reported acting like scientists. Similar to thinking like a scientist, the pre-test for acting like a scientist represented the most important predictor for this outcome at the end of the course. Self-rated scientific ability was the second most important predictor of acting like a scientist. Neither SAT scores nor high school biology grade had a direct effect on students' frequency of acting like a scientist by the end of their introductory STEM course. Students who felt more confident in their ability to learn reported acting like a scientist significantly more frequently than their peers who had less learning confidence.

Other direct effects on acting like a scientist included attending review sessions and feeling well-prepared for the course. As with thinking like a scientist, students who more frequently attended review sessions reported acting like a scientist significantly more often than their peers who attended review sessions less often. Additionally, feeling better prepared for their introductory STEM courses significantly and positively predicted students' scores on acting like a scientist at the end of the course. Being in a course that encouraged students to apply concepts in new situations also corresponded with significantly higher scores on acting like a scientist. Interestingly, students who reported having changed their study habits during the course reported acting like a scientist significantly more often than their peers who did not change their study habits.

In the focus groups, students talked extensively about lab, which is one setting where they are given the opportunity to conduct experiments and apply concepts, thus increasing their ability to act like scientists.

> Well, like how the labs really supplement the class, like they really make you think about the main concepts, about like how you would apply it to like real life or what you would actually do that shows this process of whatever. The really helps you kind of think about it other than just like bullet points on a piece of paper, so that really helps. (Marissa, Southeastern Private Master's College)

Not all students felt the way Marissa did, however. Some expressed concern that the labs did not always align with the courses and concepts they were learning. One campus, Southwestern Private Research University, had tried incorporating virtual labs, which many students said were not helpful because clicking buttons was not the same as actually conducting experiments in the lab and using problem-solving skills for troubleshooting. According to students, the lab can be a powerful place for learning how to act like a scientist, as long as the lab aligns well with the attached course and gives students the opportunity to "do" science.

Finally, one faculty-level variable, being in a classroom where faculty tried to dispel notions of competition, significantly predicted acting like a scientist more often. By dispelling notions of competition, faculty may have created a more welcoming classroom climate where students worked more collaboratively in a variety of activities, including experiments and looking up scientific journal articles. As indicated by this quote by Naomi, a student at Southeastern Public Master's College, having a faculty member assign students to a study group can be beneficial when preparing for exams while dispelling notions of competition.

> Well, I actually have to go to [a study group] after this and we, what we do is we'll, we were assigned groups before the exams we have. Like I have a biology exam on Friday.

> The group that I have, we've already established [that] we come together the Tuesday before the exam. So since this is the Tuesday before the exam on Friday, we always come together at 6:00 or 6:30, and we review the chapters, and our professor is there if we have last-minute questions with him or whatever. But I mean, to me, it's helpful because it kind of, we get to comb through what we don't really understand, and he's there to answer questions. But I think that's the one thing that is helpful and we realize once we're in that group if we don't know something, we know what we need to study, so.

Scholars consistently have demonstrated that competitive classroom climates have negative effects on students' learning and performance. For example, Walberg (1979) examined the effect of classroom climate on student achievement and retention and found that students who experienced more competitive classroom contexts tended to have higher rates of failure and lower levels of self-confidence than did their peers in more cooperative environments. Thus, by dispelling notions of competition, faculty may be constructing a more welcoming learning environment for talent development.

Discussion

Although students' grades in introductory courses may be useful for sorting students, they do not seem to be useful for capturing gains in dispositions for scientific work. As we expected, students' course grades were in large part predicted by high school preparation (SAT scores, high school biology grade), which means that "success" in introductory courses was more related to previous preparation than to science skills developed in these courses. Students who do not earn top grades do not necessarily lack the skills needed to be good scientists; they may simply lack the prior preparation or study skills needed to perform well in lecture-based classes that reward cramming and rote memorization for exams. Our findings suggest that to keep talented students in science majors, we need to broaden performance criteria and assessment techniques. Grades alone will not identify the nascent scientific talent that exists among college students.

Many NSF-funded projects are specifically devoted to interventions that are designed to improve the teaching and learning of science, yet there remains resistance to change. Our study suggests that science faculty must confront the questions of whether we can afford to cram content into students at the expense of the development of scientific skills and thinking, and whether we can continue to let grading practices reflect previous preparation rather than actual learning in the classroom. With increased interest in STEM among entering students (Higher Education Research Institute, 2010), the U.S. is at a critical crossroads in terms of its opportunity to improve the production of science degrees. In order to move forward most productively, faculty must reexamine how they assess learning while emphasizing alterative ways to identify and develop students' scientific talent in introductory STEM courses. Even faculty members in our study who said that grades are not the best predictors of talent continue to use multiple-choice tests in their courses, despite their flaws.

Currently, the vast majority of students are still in large lecture venues in introductory science and mathematics courses, and this is not likely to change. However, this does not mean that faculty cannot change the way that they teach students and develop scientific talent. Our findings suggest that allowing students to apply concepts can help them to think like scientists. Additionally, engaging with students in smaller settings (such as labs) can encourage students to act like scientists while fostering collaborative learning among students can increase their ability to succeed in these courses. Further research, however, is needed to better understand the impact of more varied and engaging pedagogies used by faculty in science.

If introductory science curriculum continues to emphasize only the transmission of content knowledge at the expense of more general higher-order thinking skills, we risk losing a significant number of future independent thinkers. Instead, if introductory courses can instill in students and reward them for mastery of critical thinking skills, we have the opportunity to develop young scientists equipped not only to master scientific concepts and knowledge but also to critique pre-existing knowledge. Investments made in these areas are necessary to open the valve in the pipeline that is preventing the movement of current students past introductory science coursework.

References

Astin, A. W. (1993). *What matters in college? Four critical years revisited*. San Francisco, CA: Jossey-Bass.

Bentler, P. M. (2006). *EQS 6 structural equations program manual.* Encino, CA: Multivariate Software, Inc.

Bentler, P. M. & Wu, E. J. C. (2002). *EQS 6 for Windows user's guide.* Encino, CA: Multivariate Software, Inc.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press

Chang, M. J., Cerna, O., Han, J., & Sáenz, V. (2008). The contradictory roles of institutional status in retaining underrepresented minorities in biomedical and behavioral science majors. *The Review of Higher Education*, *31*(4), 433-464.

Conley, D. T. (2005). *College knowledge: What it really takes for students to succeed and what we can do to get them ready.* San Francisco: Jossey-Bass.

Creswell, J. W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative approaches to research.* 2nd ed. Upper Saddle River, NJ: Merrill/Pearson Education.

Creswell, J.W. (2007).*Qualitative inquiry and research design: Choosing among five approaches.* Thousand Oaks, CA: Sage.

Gainen, J. (1995). Barriers to success in quantitative gatekeeper courses. In J. Gainen & E. W. Willemsen, (Eds.), *Fostering student success in quantitative gateway courses* .New Directions for Teaching and Learning, 61. San Francisco: Jossey-Bass.

Gasiewski, J.A., Eagan, M.K., Garcia, G.A., Hurtado, S., & Chang, M.J. (2012). From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses? *Research in Higher Education, 53*(2), 229-261.

Hagedorn, L. S., Siadat, M. V., Fogel, S. F., Nora, A., & Pascarella, E. T. (1999). Success in college mathematics: Comparisons between remedial and nonremedial first-year college students. *Research in Higher Education, 40*(3), 261-284.

Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., et al. (2004). Policy Forum: Scientific teaching. *Science, 304*(5670), 521-522.

Higher Education Research Institute (2010). *Degrees of Success: Bachelor's Degree Completion Rates among Initial STEM Majors.* Retrieved from http://www.heri.ucla.edu/nih/HERI_ResearchBrief_OL_2010_STEM.pdf

Hurd, P. D. (1998). Scientific literacy: New minds for a changing world. *Science Education, 82*(3), 407-416.

Kamii, M. (1990). Opening the algebra gate: Removing obstacles to success in college preparatory mathematics courses. *Journal of Negro Education, 59*(3), 392-406.

Labov, J. B. (2004). From the National Academies: The challenges and opportunities for improving undergraduate science education through introductory courses. *Cell Biology Education, 3*(4), 212-214.

Laird, T. F. N., Engberg, M. E., & Hurtado, S. (2005). Modeling accentuation effects:

Enrolling in a diversity course and the importance of social action engagement. *The Journal of Higher Education, 76*(4), 448-476.
Maxwell, J. (2005). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage Publications.
Merriam, S. (1998). *Qualitative research and case study applications in education.* San Francisco, CA: Jossey-Bass Publishers.
National Center for Education Statistics. (2000). *Entry and persistence of women and minorities in college science and engineering, NCES 2000-601.* Washington, DC: U.S. Department of Education.
Payzant, T. W., and Wolf, D. P. (1993). Piloting pacesetter: Helping at-risk students meet high standards. *Educational Leadership* 50(5): 42-45.
Raykov, T., Tomer, A., & Nesselroade, J. R. (1991). Reporting structural equation modeling results in *Psychology and Aging*: Some proposed guidelines. *Psychology and Aging, 6*(4), 499-503.
Seymour, E. (2001). Tracking the processes of change in US undergraduate education in science, mathematics, engineering, and technology. *Science Education, 86*(1), 79-105.
Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.
Sharkness, J., Eagan, M.K., Figueroa, T., & Hurtado, S. (2011, May). Identifying STEM Achievement: Differentiating and Predicting GPA and Gains in Scientific Competencies Among STEM Students. Paper presented at the Annual Forum of the Association for Institutional Research. Toronto, Ontario, Canada.
Teddlie, C. & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Los Angeles, CA. Sage Publications, Inc.
Tobias, S. (1990). Stemming the science shortfall at college. In S. Tobias (Ed.), *They're not dumb, they're different.* Tucson, AZ: Research Corporation.
Waits, B. K., and Demana, F. (1988). Relationship between mathematics skills of entering students and their success in college. *School Counselor*35(4): 307-310.
Walberg, H. J. (1979). *Educational environments and effects: Evaluation, policy, and productivity.* Berkeley, CA: McCutchan.
Williams, W. M., Papierno, P. B., Makel, M. C., & Ceci, S. J. (2004). Thinking like a scientist about real-world problems: The Cornell Institute for Research on Children science education program. *Journal of Applied Developmental Psychology, 25*(1), 107-126.

Appendix

Variable	Coding
Outcome Variables	
Thinking like a scientist	Construct composed of nine items (shown in Table 1)
Acting like a scientist	Construct composed of five items (shown in Table 1)
Final course grade	Continuous, range 0.0-4.0
Course-Level Variables	
Faculty agreement: Encouraged collaboration among students	1=strongly disagree to 4=strongly agree
Faculty agreement: In my classroom, there is no such thing as a question that is too elementary	1=strongly disagree to 4=strongly agree
Faculty agreement: I try to dispel notions of competition	1=strongly disagree to 4=strongly agree
Frequency of use: Essay exam questions	1=0 to 7=6 or more
Student-Level Variables	
Sex: Female	0=No, 1=Yes
Race: White (reference group: non-White)	0=No, 1=Yes
SAT composite score	Continuous, range 400-1600
HS biology grade	0=F to 4=A
Self-rating: Time management	1=Lowest 10% to 5=Highest 10%
Thinking like a scientist pre-test	Construct composed of nine items (shown in Table 1)
Acting like a scientist	Construct composed of five items (shown in Table 1)
Agreement: I considered dropping this course during the term	1=strongly disagree to 4=strongly agree
Agreement: I was well-prepared for this course's difficulty	1=strongly disagree to 4=strongly agree
Agreement: I felt my hard work was reflected in my grades	1=strongly disagree to 4=strongly agree
Agreement: This course emphasized applying concepts to new situations	1=strongly disagree to 4=strongly agree
Frequency: Attended review sessions	1=Never to 5=Very often
Frequency: Had to change my study habits during the term	1=Never to 5=Very often
Frequency: Felt confident in my ability to learn	1=Never to 5=Very often

Table 1

Factor Loadings for the latent constructs in the model

	Pre-Test	Post-Test
Thinking Like a Scientist		
Make connections between different areas of science and math	0.67	0.70
Make sense of scientific concepts	0.71	0.72
Identify what is known about a problem	0.63	0.63
Ask relevant questions	0.60	0.64
Draw a picture to represent a problem or concept	0.46	0.51
Make predictions based on existing knowledge	0.69	0.79
Come up with solutions to problems and explain them to others	0.67	0.72
Investigate alternative solutions to a problem	0.67	0.68
Translate scientific terminology into non-scientific language	0.57	0.62
Acting Like a Scientist		
Relate scientific concepts to real-world problems	0.71	0.75
Synthesize several sources of information	0.70	0.70
Conduct an experiment	0.54	0.54
Look up scientific research articles and resources	0.59	0.57
Memorize large quantities of information	0.41	0.44

Note: Measurement model fit indices: CFI=0.96 and RMSEA = 0.03.

Table 2
Direct Effects

	b	B	SE	Sig.
Thinking like a scientist (Post-Test)				
Student Level				
Acting like a scientist - pre-test	0.12	0.11	0.02	***
Thinking like a scientist - pre-test	0.37	0.38	0.03	***
I feel confident in my ability to learn	0.21	0.22	0.02	***
This course emphasizes applying concepts to new situations	0.09	0.07	0.02	***
I considered dropping this course during the term	-0.03	-0.04	0.02	*
I was well-prepared for this course's difficulty	0.07	0.06	0.03	**
Race: White	0.07	0.04	0.02	**
Frequency: Attended review sessions	0.07	0.10	0.01	***
Classroom Level				
In my classroom, there is no such thing as a question that is toc	0.03	0.57	0.02	*
Acting Like a Scientist (Post-Test)				
Student Level				
Acting like a scientist - pre-test	0.39	0.41	0.03	***
Thinking like a scientist - pre-test	0.10	0.11	0.02	***
I feel confident in my ability to learn	0.14	0.16	0.02	***
This course emphasizes applying concepts to new situations	0.12	0.11	0.02	***
I was well-prepared for this course's difficulty	0.07	0.06	0.02	**
I had to change my study habits during the term	0.03	0.04	0.01	**
Frequency: Attended review sessions	0.05	0.08	0.01	***
Classroom Level				
I try to dispel perceptions of competition	0.07	0.59	0.02	*
Final Course Grade				
Student Level				
I feel confident in my ability to learn	0.07	0.06	0.02	***
Composite SAT	0.00	0.19	0.00	***
HS Biology grade	0.30	0.15	0.05	***
I felt my hard work was reflected in my grades	0.18	0.14	0.03	***
I considered dropping this course during the term	-0.27	-0.26	0.03	***
I was well-prepared for this course's difficulty	0.13	0.09	0.03	***
Self-rated time management ability (pre-test)	0.14	0.13	0.02	***
I had to change my study habits during the term	-0.12	-0.14	0.02	***
Classroom Level				
Encouraged collaboration among students	0.09	0.25	0.05	
Used essay exams	-0.05	-0.39	0.02	**

Table 2 (con't)

	b	B	SE	Sig.
Thinking Like a Scientist Pre-test				
Composite SAT	0.00	0.07	0.00	***
Race: White	0.10	0.05	0.03	***
SAT Composite score				
Race: White	56.70	0.18	11.86	***
Gender: Female	-31.09	-0.10	5.36	***

Note: Fit indices – CFI = 0.94, RMSEA = 0.04; R^2 (Acting like a scientist) = 30.6%; R^2 (Thinking like a scientist) =33.2%; R^2 (Final grade) = 33.2%.