

## **Spiders or Butterflies? Despite Student Preference, Gender-Biased Lesson Models. Do Not Impact Interest, Attitude, and Learning in Biology**

**Amy Buxton**

Biology, Brigham Young University

**Jamie Jensen**

Biology, Brigham Young University

**Geoff Wright**

Technology and Engineering Studies  
Brigham Young University

**Seth Bybee**

Biology, Brigham Young University

**Andrea Phillips**

Biology, Brigham Young University

**Tanner Phillips**

Biology, Brigham Young University

**Michael Steadman**

Biology, Brigham Young University

### **ABSTRACT**

**Background:** Educational research often emphasizes the prevalent gender gap between males and females in science, technology, engineering, and mathematics (STEM) fields. In our study, we took a new approach on gender research by specifically considering whether there is a gender-oriented preference in curricular models (the specific lesson examples and content used to teach a broader biology topic, e.g., dust mites as a model of symbiosis) implemented to teach biology and how these models affect student interest, attitude, and learning. We sampled kindergarten through sixth grade students to determine whether a gender-oriented preference concerning lesson models exists and when that preference is most prevalent. We then designed active-learning curricula surrounding the models showing the largest gender preference and measured whether lesson model or presenter gender impacted student interest, attitude, and learning.

**Results: Our findings show that students do indeed indicate a preference to learn using their own gender-oriented lesson models from kindergarten through sixth grade, but that the lesson model and presenter gender do not impact student interest, attitude, or learning during an active learning biology presentation.**

**Conclusions: Our findings suggest that teachers as early as kindergarten should be aware and sensitive to the gender-based preferences for models used in teaching science that may exist within their classrooms and opt to alternate between male- and female-oriented lesson models to create a more inclusive classroom and to encourage especially females to pursue science. However, we offer strong advice to teachers to implement active-learning lessons as this may be the key to eliminating such gendered effects, as shown by our research.**

**Keywords:** gender-based preference, lesson models, biology education, elementary school

## INTRODUCTION

### Gender Gaps

Research has shown a prevalent gender gap in science, technology, engineering, and mathematics (STEM) fields. In the US workplace, men outnumbered women (71% vs. 29% overall) in all science and engineering fields in 2010 (NSF 2019). In academics, the gender gap is also prevalent with girls underperforming boys on biology, chemistry, statistics, physics, and calculus AP exams (College Board 2019), on standardized science tests (although the disparity decreases from 4<sup>th</sup> to 8<sup>th</sup> grade; NAEP 2019), and on the Trends in International Mathematics and Science Study (TIMSS; Provasnik, Kastberg, Ferraro, Lemanski, Roey, and Jenkins 2012). This pattern does not hold true in all countries. Researchers found that fourth through eighth grade female students in Turkey consistently had higher science success than male students, the difference becoming statistically significant as grade level increased (Bursal 2013). Another study suggested that females in Turkey had more positive attitudes towards science than males (Mihladiz, Duran, and Dogan 2011). In fact, international data from the Organisation for Economic Co-operation and Development (OECD) Programme for International Student Assessment (PISA 2019) indicate that across all OECD countries, there is no measurable gender gap in science. This suggests that there are no innate differences in ability that would affect females' ability to excel in STEM, but rather that something in the environment may be deterring females from full participation.

### Gender Stereotypes

Researchers suggest that the most prominent factors promoting the apparent gap in education are the academic environment and perceived gender stereotypes (Hall and Sandler 1982; Robelen 2012). While much of the research on gender stereotypes within education focuses on high school or college level education, some research shows that stereotypes begin at a young age (e.g., Cvencek, Meltzoff, and Greenward 2011; del Rio and Strasser 2013). Children as early as age five perceive math and science as more fit for males (del Rio and Strasser 2013; Cvencek et al. 2011; Farenga and Joyce 1999). In fact, these perceived stereotypes can influence occupational interest and these interests are tied to career choice in adulthood (Auger 2005; Trice 1991; Trice and McClellan 1993; Watson and McMahon 2005). Another study showed no significant gender gap in science interests

in first through third graders but found the gender gap to increase twenty times by the 10th to 12th grade (Baram-Tsabari and Yarden 2011). It remains unclear as to when exactly the gender preferences in biology begins. Part of our study aimed to fill this gap with reference to lesson models used to teach biological topics.

### **Influences on Gender Stereotypes and Performance**

Research has shown that implicit stereotypes exist in elementary school (Nosek, Smyth, Sriram, Lindner, Devos, Ayala et al. 2009), and that perceived gender stereotypes can affect actual school performance (Hall and Sandler 1982; Robelen 2012). Some have shown these stereotypes to be influenced by perceptions of gender expertise, especially in female students (e.g. Lockwood 2006; Moè and Pazzaglia 2006; Neuburger, Jansen, Heil, & Quaiser-Pohl 2012). Several demographic factors have also been implicated in perceived gender stereotypes and/or overall academic performance, such as the education level of parents (Mavrikaki, Koumparou, Kyriakoudi, Papacharalampous, and Trimandili 2012), household size, and parent income (Klebanov, Brooks-Gunn, and Duncan 1994). While these demographic factors are not the focus of our research, we acknowledge that they appear to play a role in academia and may very well differentially impact male and female students.

### **Gender-Preferred Lesson Models**

Many studies have found a gender-based preference when it comes to specific areas of science, e.g., chemistry, physics, biology, or engineering (e.g., Baram-Tsabari and Yarden 2008; Barmby and Defty 2006; Jones, Howe, and Rua 2000). However, little has been done to analyze the effects of specific curricular materials on student learning within each of these subjects. In our study, we focused on the curricular models used to teach each topic, specifically within the subject area of biology. We define “model” as the specific lesson examples and content used to teach a broader biology topic, such as male peacocks showing off their feathers as a model to teach the topic of sexual selection. In this way, we use the term “model” not as an explanatory framework, but as an ideal example, like a “model” organism or a “model” behavior. While a problem with overusing male-based models in education has been observed (Riddell 1989), our research represents the first systematic study of male and female preference in models to teach biology.

## **RESULTS**

### **Part 1**

A Mann-Whitney  $U$  showed that on all questions combined, males select the MIM more often than females select the MIM ( $M_{\text{females}} = 36.2\%$ ,  $N=271$ ;  $M_{\text{males}} = 65.6\%$ ,  $N=248$ ;  $U = 55,515.5$ ,  $p < .001$ ). Using student gender as a predictor for selection of the MIM, a logistic regression showed statistical significance for every question ( $p < .05$ ). In addition, a significant difference was found in every grade between males and females (see Table 2 and Figure 1) indicating that a gender-based preference is present as early as kindergarten and persists through sixth grade. The largest gender gap is in third grade (43.5%).

The four model pairs that showed the greatest gender-based preference were questions 3, 12, 15, and 16. These pairs were ladybugs versus termites (science as a process), flowers versus animal skulls (evolution), flamingos versus eagles (relationship of structure to function), and butterflies versus spiders (diversity), respectively. The percent of male and female students selecting the MIM for each of the four questions is depicted in Figure 2.

## **Part II**

### ***Interest***

An analysis of variance (ANOVA) on quantified observations showed that the model used and the presenter gender did not significantly impact the average number of comments per male student or the average number of comments per female student [Females:  $F(1) = 0.34, p = .57$ ; Males:  $F(1) = .14, p = .72$ ]. However, the lesson model had a significant main effect on the proportion of female comments to male comments within the classroom ( $F(1) = 6.78, p = .02$ ), meaning that when a FIM was used, females made almost equal the number of comments as the males (.89 female comments per male comment); whereas, when a MIM was used, females made just over half the number of comments as the males (.64 female comments per male comment), regardless of the gender of the presenter (see Table 3).

### ***Attitude***

A paired-samples t-test showed that student attitude improved overall between the pre- and post-questionnaires (talent: change of .37 points,  $p = .001, d = .33$ ; career: change of .31 points,  $p = .03, d = .22$ ;  $N=91$ ; see Table 4). However, an ANCOVA analysis showed that neither lesson model nor presenter gender differentially impacted change in attitude for either of the questions (Talent:  $F(1) = 1.35, p = .25$ ; Career:  $F(1) = 3.47, p = .07$ ).

### ***Learning***

To measure learning, we considered both the total scores (including all three questions) and the multiple-choice scores alone from the questionnaires. Paired samples t-tests showed there was overall improvement in the multiple-choice score for three of the four questionnaires (see Table 5). Total scores showed no significant differences (Questionnaire 1:  $t(86) = .93, p = .36$ ; Questionnaire 2:  $t(92) = .59, p = .56$ ; Questionnaire 3:  $t(60) = .95, p = .35$ ; Questionnaire 4:  $t(81) = .53, p = .60$ ). However, ANCOVA analyses showed that neither lesson model nor presenter gender differentially impacted learning on any of the questionnaires, all  $p > 0.05$ , see Figure 3.

## **DISCUSSION**

### **Part 1**

Our findings suggest that a gender-based preference in lesson models can be seen as early as kindergarten and that it persists through at least sixth grade with females preferring female-oriented models and males preferring male-oriented models. These results are in line with other studies that suggest gender stereotypes in academics begin at a young age (Cvencek et al., 2011; del Rio and Strasser 2013; Farenga and Joyce 1999), but they counter at least one study that found no statistical significance in gender bias in science interests in first to third graders (Baram-Tsabari and Yarden 2011). Renninger (2000) found that these preferences that lead to sustained interest can emerge as early as pre-K (age 4-5).

Our study shows that there is large gender preference in females in Kindergarten through 3<sup>rd</sup> grade (with nearly 70% or more of females choosing the female-oriented model). Whereas, in 4<sup>th</sup> and 5<sup>th</sup> grade (ages 9 – 11), female students converge on a nearly 50-50 split between the female-oriented model and the male-oriented model, suggesting less of a gender-based preference. Males, on the other hand, have the lowest preference in Kindergarten, having a nearly 50-50 split in model choice, but sustain a male-oriented bias from 1<sup>st</sup> grade on (ages 6-12). For males, this finding supports the literature that suggests that children become more aware of gendered stereotypes as age increases,

especially among boys (Kurtz-Costes, Copping, Rowley, and Kinlaw 2014). However, the female trend is more peculiar and suggests that either girls between the ages of 9 and 11 tend to see objects as being less gender-specific, or they actively seek-out what they deem as a 'male-oriented' choice. In one study of middle and high schoolers, Desy, Peterson, and Brockman (2011) found no differences in interest by gender among younger children, with differences only reaching statistical significance in high school. However, these children were all older than the present study. In addition, in our study, it appears that in sixth grade the gap begins to appear again. This may in fact correspond with the advent of puberty in girls at this age, causing greater awareness and desire to conform to a given gender stereotype (Kessels 2005). The diminishing gap between ages 9 and 11 is a point for future study.

Four questions showed the largest gaps in gender preference: ladybugs versus termites (science as a process), flowers versus animal skulls (evolution), flamingos versus eagles (relationship of structure to function), and butterflies versus spiders (diversity). Two of these four questions contrasted two arthropods, one being flashier and more colorful and one being less flashy and more dangerous. Males gravitated toward the more dangerous arthropods (termites and spider), while girls chose the colorful species (ladybugs and butterflies). Differences in preferences for biology content has been documented in many previous studies, with boys gravitating toward more graphic (e.g., autopsies) or dangerous organisms (e.g., dinosaurs) and girls preferring plants or animal care (see Cipcova, Karolcik, Dudova, and Nagyova 2018, for a review). In a third pair, boys preferred animal skulls, potentially corresponding to a preference for autopsies or even dinosaurs as they look similar to fossils (Dawson 2000); whereas, girls preferred flowers, similar to a previously documented preference for botany (Hong, Shim, and Chang 1998). In the last pair, girls preferred flamingos while boys preferred eagles. However, preference for eagles, in general, dramatically decline with age showing less than 50% of boys choosing eagles in 5<sup>th</sup> grade, and only 10% of girls choosing eagles in 6<sup>th</sup> grade. Given that these organisms are both birds and the topic related to animal form and function, it's difficult to determine why such gender disparity exists, perhaps except to suggest the color pink may have been a determining factor in early grades. The decreasing preference for eagles with age remains an enigma. However, in all four of these examples, we still see the distinctive narrowing of the gap in 4<sup>th</sup> and 5<sup>th</sup> grades, mostly due to a higher preference for male-oriented models amongst girls.

Regardless of the individual patterns, our data clearly show a distinct preference among each gender for their gender-oriented lesson model. This differential preference is present in as early as Kindergarten, although it likely begins earlier (Renninger 2000), and persists through sixth grade. The literature suggests that it persists and intensifies in middle school and beyond (e.g., Desy, Peterson, & Brockman 2011).

## Part 2

Our overall findings were contrary to our original hypotheses. Although male and female students select different teaching models that they would prefer to use (as shown in Part 1 of our study), the model used and the presenter gender do not appear to impact the interest, attitude, or learning of male or female students. Interestingly, students were equally engaged in the presentation, showing a high level of participation. This suggests the need for an alternative hypothesis for the equal success of each model.

Our alternative hypothesis is that teaching methods play a larger role than model or presenter gender in third grade student interest, attitude, and learning. Because each of our lessons was designed to be inquiry-based and active learning, students were actively involved in the learning process. While male and female students may indicate a different preference for models, perhaps an inquiry-based, active learning lesson engages students of both genders, regardless of the model used or the presenter's gender. This should be a key area of future studies. Engaging student interest for both genders is particularly important considering that elementary school students' biology interests have declined over the course of a generation, from 1980 to 2011 (Randler, Osti, & Hummel, 2012).

### ***Interest***

The finding that neither model gender or presenter gender impacted student interest is in contrast to what we anticipated, as we thought that student interest would increase if we selected matched models. Our initial hypothesis aligned with a study showing that female students were less interested in computer science than male students when the classroom was filled with stereotypically male objects, such as *Star Trek* posters and video game paraphernalia, but they were equally interested when these objects were removed (Cheryan, Plaut, Davies, & Steele, 2009). A significant finding however was that the ratio of female to male comments per student increased when a female-interest model was used ( $F(1) = 6.78, p = .02$ ). In other words, when we taught using a female-interest model, female hand-raising would increase as compared to male hand raising. This could have implications for teachers struggling with a classroom dominated by male comments. In order to balance student participation, those teachers might consider implementing more female-interest models into their curriculum.

### ***Attitude***

Despite no impact of the model or presenter gender, student attitude (based on student surveys about their perceived talent in biology or career aspirations in science) improved from pre- to post-presentation. We proposed that choosing appropriate models for each gender could help students find more value in biology, and thus increase their career aspirations; this was based on previous literature on child career development indicating that occupational gender stereotyping is seen in elementary school and that it influences children's career aspirations (e.g., Auger et al., 2005; Watson & McMahon, 2005). In fact, Gottfredson's influential theory on career development suggested that younger children's career aspirations are especially likely to be influenced by sex-typing (Gottfredson, 1981). This theory has received further support in other studies (see Watson & McMahon, 2005, for a review). In addition, students' occupational choice is partially dependent on the value they give a subject (Koul, Lerdpornkulrat, & Chantara, 2011). We likewise proposed that lesson model and presenter gender could be key to overcoming the implicit stereotypes underlined by Nosek et al. (2009), thereby improving student attitude. Exposing children to a non-traditional worker (e.g., a female scientist) in early elementary school has resulted in reduced perception of occupational gender stereotype in past studies (Bailey & Nihlen, 1990; Bigler & Liben, 1990). Based on our results, however, it seems that lesson model and presenter gender are not what most impacts career aspirations or perceived talent in biology, at least at this young age. This finding is similar to what we observed for student interest, further advocating the need of our alternative hypothesis: that an active learning, inquiry-based lesson may have the ability to overcome these perceptions and improve student attitude, regardless of the gender of the presenter or model used to teach.

### ***Learning***

Although learning was not differentially impacted by model or presenter gender, students showed improvement on three of the four questionnaires when we looked at the easily quantifiable multiple-choice responses alone. This is interesting to consider in light of the work by del Rio and Strasser (2013) that found a gender stereotype in mathematics to exist in students at an age when there was no difference in performance suggesting that students may show a gender-based preference concerning what topics or models are appropriate or preferred for each gender, without those preferences actually impacting performance. These findings are parallel to what we saw with interest and attitude, further underlining the potential importance of teaching method.

It is likewise important to note that improvement was marginal and was in fact not demonstrated on the multiple-choice question for one questionnaire and on the total score for all four questionnaires. This brings to mind at least two possible explanations. First, students may have been unaccustomed to the question format used in the study. In support of this possibility, our questionnaire 3 multiple-choice results showed the most improvement out of all questionnaire multiple-choice questions. This particular question is the only one that was written at the “remember” level of Bloom’s taxonomy (Bloom, 1984). Because third graders show more improvement on this question, it seems reasonable to consider whether this is the question format to which they are most accustomed. In asking for question samples from the teachers before writing the questionnaires, we noticed a trend toward all remember-level questions. It would be interesting to further investigate whether third grade students are ever given science questions utilizing upper levels of Bloom’s taxonomy in science.

Second, our findings suggest that we may be trying to test student learning on a level they have not yet reached. It is possible that many of the reasoning patterns and the conceptual understanding we were trying to teach and assess may have been beyond their developmental level or ability to accomplish, even with appropriate scaffolding (Vygotsky, 1978). Interestingly, our third topic (diversity) dealt with the simple, concrete concept of conservation and was assessed in a more concrete fashion. As noted above, students showed the most improvement on the multiple-choice question for this topic, lending further support to this hypothesis. Future work might include investigating gender-based preferences in learning at a higher grade level where students are at a higher developmental level. It seems likely, however, that the patterns in gender-based preference and learning that we observed were an accurate depiction of reality, as they parallel the patterns we observed in interest and attitude.

### ***Presenter Gender***

We were surprised to find that presenter gender had no significant impact on the number or proportion of student comments, improvement in student attitude, or improvement in student questionnaire scores. Our findings countered what we had expected based on several prior studies that found that participant scores on mental-rotation performance tasks depended on what they were told about their own or the other gender’s aptitude for the task (Moè & Pazzaglia, 2006; Neuburger, Jansen, Heil, & Quaiser-Pohl, 2012). Based on these studies, we had predicted that if males and females often perform better when they believe their gender is good at a particular task, then presenter gender might impact student performance as the students would see someone of their own gender with an aptitude for biology. In addition, research suggesting that same-gender role models benefit female college students seemed to support our prediction (Lockwood, 2006).

Despite these findings and our initial hypothesis, we found that student interest, attitude, and learning do not appear to be impacted by presenter gender. This could have implications for schools concerned with an unequal number of male and female teachers. If teacher gender, like presenter gender, is found to have little to no impact on student participation, this may help to mitigate this concern. In fact, one empirical study conducted at an English primary school already suggests such a trend, finding that matching teacher and student gender had no impact for male or female student achievement or attitude toward school (Carrington, Tymms, & Merrell, 2008). Together with these findings, our study suggests that having equally represented male and female presenters or teachers in the elementary school does not inherently improve student interest, attitude, or learning, at least in biology.

### **Implications**

Does it matter that students have different preferences for the models used in a science classroom and that these preferences seem to revolve around gender? Hidi and Renninger (2006) suggested a four-phase model of interest development. The first two phases, triggered situational interest and sustained situational interest, are almost exclusively influenced by the external environment, i.e., to what students are being exposed in the classroom. Student situational interest is heavily dependent on a student's positive feelings toward the topic. Thus, a more gender-normative topic may spark higher situational interest. Situational interest then can lead to emerging or sustained individual interest, a state at which a student is more likely to reengage voluntarily with the topic. Interest has been shown definitively to promote academic behaviors that lead to better achievement in science (Hidi & Renninger 2006). Liebham, Alexander, and Johnson (2013) found that this early interest in science influences later science achievement and self-concept in science for girls, but not boys. Another study in Thailand found that students' choice of occupation depends in part on the value they give to a subject early on (Koul, Lerdpornkulrat, and Chantara 2011). Early interest in science may indeed be a potential key in closing the gender gap in science achievement and retention for females. Thus, the lesson model chosen may have a profound effect on science identity, science achievement, and later persistence in a science career.

Nosek and others (2009) suggest that attempts to boost women's participation and achievement in mathematics and science need to overcome the implicit stereotypes in people's minds. By consistently choosing one gender-oriented lesson model over another (e.g., male-oriented models over female-oriented models) it seems quite possible that we are currently driving female students away from STEM subjects and thereby supporting gender stereotypes by the way we teach. For example, in a study in computer science, researchers changed the posters displayed in the classroom from those typical of male interests (e.g., Star Trek or video games) to those not associated with any stereotype (e.g., nature posters). They found that this change was sufficient to boost undergraduate female students' interest in computer science to that of their male counterparts (Cheryan, Plaut, Davies, and Steele 2009). If something as simple as classroom environment can perpetuate gender stereotypes, it seems likely that the models we use within the classroom may also influence student interest. We would suggest, however, that implementing a more active-learning style of teaching may mitigate these effects. Many studies have been done on the gendered portrayal of people in science textbooks and curricular material (e.g., World Science Report 1996; Kahveci 2010), however, little to no work has been done in studying the gender-orientation of the lesson materials (i.e., models) used in teaching science, with one notable exception (Colette and Marjolaine 2017). Our findings suggest that teachers as early as

kindergarten should be aware and sensitive to the gender-based preferences for models used in teaching science that may exist within their classrooms and opt to alternate between male- and female-oriented lesson models to create a more inclusive classroom and to encourage especially females to pursue science. However, we offer strong advice to teachers to implement active-learning lessons as this may be the key to eliminating such gendered effects, as shown by our research.

## CONCLUSIONS

This project has revealed useful information about elementary school biology education and can also be a launching point for future research. We envision that related research will expand this study to grade levels beyond elementary school and to other STEM fields. Importantly, our study was limited to a suburban low-minority, semi-affluent population of students; thus, results may or may not be generally applicable to other populations and should be taken thoughtfully until further study can be done on alternative populations.

Notwithstanding, these results may be particularly important to consider when trying to combat the gender gap in fields where females are more drastically underrepresented (e.g., physics or engineering). While men outnumber women in STEM careers in the workplace (NSF 2018), the National Science Board (2010) suggests that since 2000, women have earned approximately half of science and engineering bachelor's degrees. In addition, recent statistics suggest that female students have equal achievement to males in science and math at the kindergarten-Grade 12 (K-12) level, but their participation in higher level courses still lags in computer science and engineering (NSF, 2018). However, there are still variations in which gender earned the most degrees in which fields at the post-secondary level. For example, while men earned more bachelor's degrees in engineering, computer sciences, and physics, women actually earned more than half of the biological science degrees (NSF 2019). Other studies likewise show a pattern of male interest in physics and female interest in biology (Baram-Tsabari and Yarden 2008; Barmby and Defty 2006; Farenga and Joyce 1999; Jones et al. 2000). This pattern suggests that future research related to our study could successfully expand this work on gender preference for lesson models to other age levels and STEM fields in hopes of discovering important implications for those groups.

## METHODS

Our study included two parts. The purpose of part 1 was to determine if and when a gender-based preference in lesson models appears among elementary students, grades kindergarten through sixth, and how pronounced this gender-based preference is. In part 2, we chose the four model pairs that showed the greatest gender-based preference and used them to create and implement lesson plans in the grade level where gender-based preference was shown to be the largest. From these lessons, we determined whether the gender bias of the model and/or the gender of the presenter impacted students' interest, attitudes, and learning.

### Part 1

#### ***Gender Preference Instrument.***

The instrument focused on lesson models that could be used to teach certain topics in biology. Topics were chosen from the Next Generation Science Standards [NGSS 2015; science as a process (weaved through the NGSS), evolution (Life Science section 4B; LS4B), animal behavior (LS2D), ecology (LS2C), relationship of structure to function (LS1A), diversity (LS4A), cells (LS1A), genetics (LS3A), and science and society (LS4D)]. We created an instrument containing 24 juxtaposed pairs

of lesson models for each topic: one model that is traditionally of stereotypical male interest (MIM), and one model that is traditionally of stereotypical female interest (FIM). For example, teaching diversity using spiders (MIM) or butterflies (FIM). Other examples include teaching evolution using animal skulls (MIM) versus flowers (FIM), the relationship of structure to function using sharks (MIM) versus dolphins (FIM), or ecology using lice parasitism (MIM) versus bird parasitism (FIM). The images chosen were similar enough in style and visual appeal to prevent superfluous preference. (See Online Resource 1.) Our initial selection of FIMs and MIMs was purely trivial and informed by popular cultural references and cultural clichés and norms, e.g., school supplies targeted at female students are more likely to portray dolphins, while the same school supplies targeted at male students are more likely to display sharks (e.g., Barthel 1988; Bretl and Cantor 1988). We did not presume to have accurate gender-oriented models a priori, but rather allowed the data to dictate after the fact whether a gender preference was actually present. Student responses dictated whether or not a gender-orientation was evident in our chosen pair.

### ***Instrument Administration***

We administered our instrument to students at two local elementary schools consisting primarily of suburban, low-minority, semi-affluent students. We emphasize that these schools were chosen based on proximity as well as a willingness to participate. Caution should be taken in generalizing findings beyond a similar student demographic. We surveyed students in each grade as follows: 33 kindergarten students, 58 first grade students, 58 second grade students, 99 third grade students, 101 fourth grade students, 82 fifth grade students, and 88 sixth grade students (a total of 271 girls and 248 boys). Students were presented with both a paper copy of the instrument as well as an image presented via PowerPoint. One pair of images at a time were shown. With the images still projected, students were asked, “If we are going to learn about a certain topic (e.g., diversity), which of the following lesson models (e.g., spiders or butterflies) would you rather use?” Students were asked not to work ahead and to stay right with the researcher’s presentation to be sure that everyone understood each question. If a student(s) did not understand a question, the researcher talked the student(s) through their question(s). This was not a common occurrence. Teachers circulated throughout the room providing help and feedback. All students were allowed ample time to make their choices and, while we acknowledge the possibility of survey fatigue, anecdotal observations indicate this was not the case as students seemed genuinely excited for the potential to choose options for learning.

### ***Data Analysis***

To determine at which grade level gender-based preference was most prevalent, we analyzed the average male and female preference data for FIMs and MIMs in each grade. Our data violated the assumption of normality (Shapiro-Wilk,  $p < .001$ ), therefore we ran non-parametric independent samples tests. To determine which models showed a significant gender-based preference, we ran a logistic regression with the independent variables being gender and grade, and the dependent variable being whether or not students chose the MIM. Statistical significance was set at .05.

## **Part 2**

### ***Presentations***

We created parallel lesson plans for the four topics showing the most gender-based preference. We used a learning cycle model, allowing students to “explore” the content before we “explained” the material (Bybee 1993). We attempted to make the lessons as parallel as possible, changing only the

model itself. In addition, both a male and female presenter (who were siblings and shared many physical characteristics as well as demeanors) taught both the MIM and FIM in different classrooms. Table 1 shows the presentation schedule. All lesson plans are outlined in the Supplemental Materials.

### ***Instruments***

To measure student interest, we recorded the classroom during each presentation. Three researchers reviewed the recordings, noting how many students of each gender made comments (or raised their hands to make comments, even if they were not called upon). This served as a proxy for student interest. Some recent studies have shown that females participate less, at least at the undergraduate level (e.g., Eddy, Brownell, & Wenderoth 2017); while other studies suggest females are actually higher achievers and therefore more likely to be comfortable raising their hands (Economist.com 2015; Gnaulati 2014). However, Shakeshaft (1995) showed that at the elementary level, girls and boys show an equal interest in science. Weighing this evidence, we chose to use hand-raising as an adequate measure for engagement. Observation protocol sheets are included in the Supplemental Materials.

To measure attitude, we included two attitude questions on the pre- and post-presentation questionnaires for the first lesson. The first of these questions (hereafter referred to as “talent”) asked, “Do you think you are good at biology?” This question required students to rate themselves on a 5-point Likert scale from ‘very’ to ‘not at all’. The second attitude question (hereafter referred to as “career”) asked, “When you grow up, would you like to have a job doing science?”, and was also graded on a 5-point scale.

To measure learning, we developed pre- and post-presentation questionnaires for each presentation’s content. Students took the pre-questionnaire within a week before we arrived, and the identical post-questionnaire within one day post instruction. The questionnaire included a single multiple-choice question followed by a free response opportunity to explain their answer. It then asked students to draw and label a picture showing a certain phenomenon, based on an instrument used by Jensen (2014). A pre- and post-questionnaire was obtained for all classes except for a missing pre-questionnaire on topic 3 for one class due to unforeseen circumstances. Questionnaires and Grading Rubrics are included in the Supplemental Materials.

We also collected demographic information about the student from parents in conjunction with obtaining permission for their child’s participation in the study. Demographics collected included mother’s education, father’s education, income level, and number of children in the family.

### ***Data Analysis.***

#### ***Interest.***

We analyzed inter-rater reliability using Pearson’s correlation coefficients between raters. Correlations between all raters were greater than 93.3% ( $p < .001$ ). We ran a series of ANOVAs on female comments, male comments, and the proportion of female and male comments in relation to model gender and presenter gender.

### **Attitude**

We ran an analysis of covariance (ANCOVA) for both talent change and career change using mother's education, father's education, income level, and number of children in the family as covariates.

### **Learning**

Rubrics were created for each item on each survey. Three independent researchers graded the questionnaires according to the rubric. After scores were assigned, researchers discussed mismatched scores until they came to full agreement.

We ran a paired samples t-test to determine whether there was overall significant improvement in the total score or multiple-choice score only from pre to post. We ran an ANCOVA on the total improvement (post-pre) for each questionnaire, looking at either student gender by model gender or student gender by presenter gender using the previously listed covariates. Students who got both the pre- and post- multiple-choice questions correct were excluded from the analyses.

### **List of Abbreviations**

**ANOVA:** Analysis of Variance

**ANCOVA:** Analysis of Covariance

**K-12:** Kindergarten-Grade 12

**LS:** Life Science

**FIM:** Female Interest Model

**MIM:** Male Interest Model

**NGSS:** Next-Generation Science Standards

**OECD:** Organisation for Economic Co-operation and Development

**PISA:** Programme for International Student Assessment

**STEM:** Science, Technology, Engineering, and Mathematics

**TIMSS:** Trends in International Mathematics and Science Study

### **References**

Auger, R. W. (2005). The development of elementary-aged children's career aspirations and expectations. *Professional School Counseling*, 8(4), 322-329.

Bailey, B. A., & Nihlen, A. S. (1990). Effect of experience with nontraditional workers on psychological and social dimensions of occupational sex-role stereotyping by elementary school children. *Psychological Reports*, 66, 1273-1282.

Baram-Tsabari, A., & Yarden, A. (2011). Quantifying the gender gap in science interests. *International Journal of Science and Mathematics Education*, 9(3), 523-550.

Baram-Tsabari, A. & Yarden, A. (2008). Girls' biology, boys' physics: evidence from free-choice science learning settings. *Research in Science & Technological Education*, 26(1), 75-92. <http://dx.doi.org/10.1080/02635140701847538>

Barmby, P., & Defty, N. (2006). Secondary school pupils' perceptions of physics. *Research in Science & Technological Education*, 24(2), 199-215. <http://dx.doi.org/10.1080/02635140600811585>

Barthel, D. (1988). *Putting on appearances*. Philadelphia: Temple University Press.

Bigler, R. S., & Liben, L. S. (1990). The role of attitudes and interventions in gender-schematic processing. *Child Development*, 61, 1440-1452.

Bloom, B. S. (1984). *Taxonomy of educational objectives*. Boston, MA: Allyn and Bacon.

- Bursal, M. (2013). Longitudinal investigation of elementary students' science academic achievement in 4-8<sup>th</sup>: Grade level and gender differences. *Educational Sciences: Theory & Practice*, 13(2), 1151-1156.
- Bretl, D. J., & Cantor, J. (1988). The portrayal of men and women in U.S. television commercials. *Sex Roles*, 18, 595-609.
- Bybee, R. (1993). An instructional model for science education. Developing Biological Literacy, Colorado Springs, CO: Biological Sciences Curriculum Studies.
- Carrington, B., Tymms, P., & Merrell, C. (2008). Role models, school improvement and the 'gender gap'—do men bring out the best in boys and women the best in girls? *British Educational Research Journal*, 34(3), 315-327. doi: 10.1080/01411920701532202
- Cheryan, S., Plaut, V. C., Davies, P. G., & Steele, C. M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology*, 97(6), 1045-1060. doi: 10.1037/a0016239
- Cipkova, E., Karolcik, S., Dudova, N., & Nagyova, S. (2018). What is the students' interest in biology after the biology curriculum modification? *The Curriculum Journal*, 29(3), 370-386.
- Colette, A., Marjolaine, C. (2017). Perception by French students of the gendered nature of material artifacts studied in technology education. *International Journal of Technology and Design Education*, 27, 1-18.
- College Board. (2019). *National Report*. Retrieved 3/18/20 from <https://research.collegeboard.org/programs/ap/data/participation/ap-2019>
- Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (2011). Math-gender stereotypes in elementary school children. *Child Development*, 82(3), 766-779. <http://dx.doi.org/10.1111/j.1467-8624.2010.01529.x>
- Dawson, C. (2000). Upper primary boys' and girls' interests in science: Have they changed since 1980? *International Journal of Science Education*, 22(6), 557-570.
- del Rio, M. F., & Strasser, K. (2013). Preschool children's beliefs about gender differences in academic skills. *Sex Roles*, 68, 231-238. <http://dx.doi.org/10.1007/s11199-012-0195-6>
- Desy, E. A., Peterson, S. A., & Brockman, V. (2011). Gender differences in science-related attitudes and interests among middle school and high school students. *Science Educator*, 20(2), 23-30.
- Economist.com (2015). The economist explains why girls do better at school than boys. Retrieved 3/18/20 from <http://www.economist.com/blogs/economist-explains/2015/03/economist-explains-3>
- Eddy, S. L., Brownell, S. E., & Wenderoth, M. P. (2017). Gender gaps in achievement and participation in multiple introductory biology classrooms. *CBE-Life Sciences Education*, 13(3), 478-493.
- Farenga, S. J., & Joyce, B. A. (1999). Intentions of young students to enroll in science courses in the future: An examination of gender differences. *Science Education*, 83(1), 55-75. [http://dx.doi.org/10.1002/\(SICI\)1098-237X\(199901\)83:1<55::AID-SCE3>3.0.CO;2-O](http://dx.doi.org/10.1002/(SICI)1098-237X(199901)83:1<55::AID-SCE3>3.0.CO;2-O)
- Gnautati, E. (2014). Why girls tend to get better grades than boys do. *The Atlantic*, retrieved 3/18/20 from <http://www.theatlantic.com/education/archive/2014/09/why-girls-get-better-grades-than-boys-do/380318/>.
- Gottfredson, L. S. (1981). Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counseling Psychology*, 28, 545-579.
- Hall, R. M., & Sandler, B. R. (1982). *The classroom climate: A chilly one for women?* Washington, DC: Project on the Status and Education of Women, Association of American Colleges.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist* 41(2), 111-127
- Hong, J. L., Shim, K. C., & Chang, N. K. (1998). A study of Korean middle school students' interests in biology and their implications for biology education. *International Journal of Science Education*, 20(8), 989-999.
- Jensen, E. (2014). Evaluating children's conservation biology learning at the zoo. *Conservation Biology*, 28(4), 1004-1011. doi: 10.1111/cobi.12263

- Buxton, A., Jensen, J., Wright, G., Bybee, S., Phillips, A., Phillips, T., & Steadman, M. (2020). Spiders or Butterflies? Despite Student Preference, Gender-Biased Lesson Models Do Not Impact Interest, Attitude, and Learning in Biology. *Advances in Social Sciences Research Journal*, 7(4) 87-101
- Jones, M. G., Howe, A., & Rua, M. J. (2000). Gender differences in students' experiences, interests, and attitudes toward science and scientists. *Science Education*, 84(2), 180-192. [http://dx.doi.org/10.1002/\(SICI\)1098-237X\(200003\)84:2<180::AID-SCE3>3.0.CO;2-X](http://dx.doi.org/10.1002/(SICI)1098-237X(200003)84:2<180::AID-SCE3>3.0.CO;2-X)
- Kahveci, A. (2010). Quantitative analysis of science and chemistry textbooks for indicators of reform: A complementary perspective. *International Journal of Science Education* 32(11), 1495-1519
- Kessels, U. (2005). Fitting into the stereotype: How gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *European Journal of Psychology of Education*, 20, 309-323.
- Klebanov, P.K., Brooks-Gunn, J., & Duncan, G.J. (1994) Does neighborhood and family poverty affect mothers' parenting, mental health, and social support? *Journal of Marriage and the Family*, 56, 441-455.
- Koul, R., Lerdpornkulrat, T., & Chantara, S. (2011). Relationship between career aspirations and measures of motivation toward biology and physics, and the influence of gender. *Journal of Science Education and Technology*, 20, 761-770. doi: 10.1007/s10956-010-9269-9
- Kurtz-Costes, B., Copping, K. E., Rowley, S. J., & Kinlaw, C. R. (2014). Gender and age differences in awareness and endorsement of gender stereotypes about academic abilities. *European Journal of Psychol Educ*, 29, 603-618.
- Leibham, M. B., Alexander, J. M., & Johnson, K. E. (2013). Science interests in preschool boys and girls: relations to later self-concept and science achievement. *Science Education*, 97(4), 574-593.
- Lockwood, P. (2006). "Someone like me can be successful": Do college students need same-gender role models? *Psychology of Women Quarterly*, 30, 36-46.
- Mavrikaki, E., Koumparou, H., Kyriakoudi, M., Papacharalampous, I., & Trimandili, M. (2012). Greek secondary school students' views about biology. *International Journal of Environmental & Science Education*, 7(2), 217-232.
- Mihladız, G., Duran, M., & Dogan, A. (2011). Examining primary school students' attitudes towards science in terms of gender, class level and income level. *Procedia Social and Behavioral Sciences*, 25, 2582-2588. doi:10.1016/j.sbspro.2011.04.150
- Moè, A., & Pazzaglia, F. (2006). Following the instructions! Effects of gender beliefs in mental rotation. *Learning and Individual Differences*, 16, 369-377. <http://dx.doi.org/10.1016/j.lindif.2007.01.002>
- NAEP, National Assessment of Educational Progress. (2019). Retrieved 3/18/20 from <https://www.nationsreportcard.gov/mathematics/nation/groups/?grade=4>
- Neuburger, S., Jansen, P., Heil, M., & Quaiser-Pohl, C. (2012). A threat in the classroom: Gender stereotype activation and mental-rotation performance in elementary-school children. *Zeitschrift fur Psychologie*, 220(2), 61-69. <http://dx.doi.org/10.1027/2151-2604/a000097>
- NGSS, Next generation science standards. (2015). Retrieved 4/5/19 from <http://www.nextgenscience.org/>
- Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A. . . Greenwald, A. G. (2009). Steele, C.M. (Ed.). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *PNAS*, 106(26), 10593-10597. [www.pnas.org/cgi/doi/10.1073/pnas.0809921106](http://www.pnas.org/cgi/doi/10.1073/pnas.0809921106)
- NSB, National Science Board. (2010). *Science and engineering indicators (National Science Foundation Report, 10-01)*. Retrieved 4/5/19 from <https://nsf.gov/statistics/seind/>
- NSF, National Science Foundation, Division of Science Resources Statistics, Arlington, VA: (2019). *Women, Minorities, and Persons with Disabilities in Science and Engineering: NSF* Retrieved 3/18/20 from <https://nces.nsf.gov/pubs/nsf19304/>
- NSF, National Science Foundation, Science and Engineering Indicators (2018). Retrieved 3/18/20 from <https://nsf.gov/statistics/2018/nsb20181/>
- PISA, OECD (2019). Programme for International Student Assessment. Retrieved 3/18/20 from <https://www.oecd.org/pisa/PISA%202018%20Insights%20and%20Interpretations%20FINAL%20PDF.pdf>
- Provasnik, S., Kastberg, D., Ferraro, D., Lemanski, N., Roey, S., & Jenkins, F. (2012). Highlights From TIMSS 2011: Mathematics and Science Achievement of U.S. Fourth- and Eighth-Grade Students in an International Context: NCES

2013-009. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Randler, C., Osti, J., & Hummel, E. (2012). Decline in interest in biology among elementary school pupils during a generation. *Eurasia Journal of Mathematics, Science & Technology*, 8(3), 201-205. doi: 10.12973/eurasia.2012.8325a

Renninger, K. A. (2000). Individual interest and its implications for understanding intrinsic motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 375-407). San Diego, CA: Academic Press.

Riddell, S. (1989). Pupils, resistance and gender codes: A study of classroom encounters. *Gender and Education*, 1(2), 183-197.

Robelen, E. W. (2012). STEM gender gap pronounced in U.S. *Education Week*, 31(36), 8-9.

Shakeshaft, C. (1995). Reforming science education to include girls. *Theory into Practice*, 34(1), 74-79.

Trice, A. D. (1991). A retrospective study of career development: 1. Relationship among first aspirations, parental occupations, and current occupations. *Psychological Reports*, 68, 287-290.

Trice, A. D., & McClellan, N. (1993). Do children's career aspirations predict adult occupations? An answer from a secondary analysis of a longitudinal study. *Psychological Reports*, 72, 368-370.

Vygotsky, L. S. (1978) *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Watson, M., & McMahon, M. (2005). Children's career development: A research review from a learning perspective. *Journal of Vocational Behavior*, 67, 119-132.

World science report-UNESCO (1996). The gender dimension of science and technology. Retrieved September 19, 2019 from <http://www.unesco.org/new/en/natural-sciences/science-technology/prospective-studies/unesco-science-report/world-science-report-1996/>