Social Psychology of Education (2026) 29:113
https://doi.org/10.1007/s11218-026-10294-y

It doesn't always have to be this way: promoting girls' interest in STEM through growth mindsets and structural explanations of gender differences

Allison Master[1] · Kahyun Lee[1] · Shaila Sharmin[1] · Daijiazi Tang[1,2] · Andrew N. Meltzoff[3,4] · Sapna Cheryan[3]

Received: 29 July 2025 / Accepted: 2 September 2026

Abstract

Gender gaps in motivation for STEM fields emerge at an early age and generally persist as children become older. Viewing these gender differences in motivation as malleable rather than fixed or inherent may reduce stereotyping and promote girls' motivation in STEM. The current paper reports findings from two preregistered experimental intervention studies. These interventions targeted girls' beliefs about stability (Study 1) and external attributions for gender differences (Study 2). Study 1 ($N = 381$) examined whether a growth mindset intervention could increase Grade 6 and 8 students' motivation for computer science. There was some evidence that girls in the intervention condition reported higher motivation on some outcomes, and exploratory findings further confirmed the importance of girls' motivational beliefs. Study 2 ($N = 220$) examined whether teaching 11- and 12-year-old girls about external, structural explanations for gender differences could influence their interest in a novel STEM activity. Exploratory results showed that structural explanations dampened gender stereotypes and also promoted belief in the malleability of STEM interest; effects on girls' motivation were not found. Beliefs that interest and abilities in STEM are malleable may be helpful but not sufficient to foster girls' own motivation in STEM.

Keywords Gender · STEM · Motivation · Growth mindsets · Structural explanations · Adolescents

✉ Allison Master
amaster@uh.edu

[1] Department of Psychological, Health, and Learning Sciences, University of Houston, Houston, TX, USA

[2] Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA

[3] Department of Psychology, University of Washington, Seattle, WA, USA

[4] Institute for Learning & Brain Sciences, University of Washington, Seattle, WA, USA

1 Introduction

Gender gaps in motivation for STEM emerge at an early age and generally persist as children get older. Girls report lower interest and self-efficacy than boys for STEM fields as early as elementary school, especially for computer science and engineering (Master et al., 2017, 2025; Nagy et al., 2010). Girls also report lower identification than boys with math and science by elementary and middle school (Cvencek et al., 2011; Vincent-Ruz & Schunn, 2018).

Debates about gender gaps in motivation and performance in STEM have centered around the extent to which these gaps are innate, biological, and fixed, versus shaped by social and cultural influences in the environment (Reges, 2018; Schmidt, 2011). One well-known example of the innate view involves James Damore's "Google Memo," which claimed, "the distribution of preferences and abilities of men and women differ in part due to biological causes and… these differences may explain why we don't see equal representation of women in tech" (Damore, 2017, p. 3). In contrast, much social science research has emphasized the role of societal stereotypes and social influences on the development of motivation and performance in STEM (Master, 2021; Master & Meltzoff, 2020; Master et al., 2021, 2025). These differing perspectives have important implications for interventions that seek to reduce gender gaps in STEM. If gender gaps are largely fixed and inherent, then interventions to improve girls' motivation in STEM may be unnecessary or unsuccessful. In contrast, if gender gaps are shaped by experiences, then well-designed interventions may be an effective way to broaden participation in STEM and promote equity. Beyond the effectiveness of interventions, whether young women themselves view these gaps as malleable versus fixed may have important implications for their personal motivation in STEM.

In the next sections, we review research on the effects of gender stereotypes on STEM motivation. We then describe two approaches for interventions that seek to boost girls' motivation and performance in STEM. Both approaches involve highlighting the malleability of STEM motivation, using growth mindsets and structural explanations to counter stereotypes by teaching students that personal characteristics and gendered patterns of interest can change. Study 1 targeted beliefs about whether the *self* can change (growth mindsets). Study 2 targeted beliefs about whether *groups* can change (structural explanations rather than essentialist explanations). The focus of this paper is not on the origins of gender differences in STEM (Wang & Degol, 2013), but instead on the negative consequences for children's educational outcomes of the *belief* that those differences are fixed and innate. Children in late elementary school and middle school are learning about academic fields and beginning to choose career paths, making these ideal stages to counter stereotypes and influence their academic motivation (Maltese & Tai, 2010).

1.1 Gender stereotypes in STEM

Stereotypes can impact students' motivation through social identity threat. Social identity threat happens when individuals from stereotyped groups feel judged, devalued, or stigmatized based on negative stereotypes about their group (Easterbrook &

Hadden, 2021; Master et al., 2026; Master & Meltzoff, 2020). These concerns about stereotypes can lead to decreased motivation and, in the case of stereotype threat, reduced performance (Murphy et al., 2024). Children and early adolescents are also vulnerable to the effects of social identity threat (Master et al., 2021; Régner et al., 2014).

Children begin to learn about and endorse STEM-gender stereotypes during elementary school (Master, 2021; Master & Meltzoff, 2020; Régner et al., 2014). Most research has focused specifically on math stereotypes, with fewer studies involving science or technology (Miller et al., 2024). Depending on the sample and method of measuring stereotypes, children appear to endorse gender stereotypes about interest in STEM by mid-elementary school (Archer et al., 2012; Cvencek et al., 2011; Farenga & Joyce, 1999; Master et al., 2021, 2025), and about ability in STEM by late elementary school or adolescence (Miller et al., 2024; Passolunghi et al., 2014). Stereotypes also differ by field. Evidence suggests that children may endorse interest and ability stereotypes about computer science and engineering earlier, by age six (Master et al., 2017, 2025).

Stereotypes can influence children's interest and motivation in academic activities. Many correlational studies have reported associations between stereotypes and various aspects of students' motivation, including: (a) girls' and women's lower participation in STEM fields (Lane et al., 2012; Miller et al., 2015; Wynn & Correll, 2017), and (b) interest in STEM (Cheryan et al., 2013a, b; Master et al., 2016, 2025; Plante et al., 2013). Several experimental studies that targeted gender stereotypes about who is interested in STEM successfully increased students' motivation, suggesting a causal link (Cheryan et al., 2011; Master et al., 2016, 2021; Murphy et al., 2007; Stout et al., 2011). For example, in one experimental study, 8- and 9-year-old children were told about two computer science activities. One activity was linked to a gender stereotype ("girls are much less interested in this activity than boys"), while the other activity was not. The gender gap in interest was only evident for the computer science activity that was linked to a gender stereotype (Master et al., 2021). This suggests that gender stereotypes can causally reduce girls' motivation for STEM, leading to the formation of gender gaps. Finding ways to counter these stereotypes may be an effective way to reduce gender gaps.

1.2 How to counter stereotypes

Stereotypes are notoriously difficult to change (Bigler, 1999; Rothbart, 2001; Steele, 2003). As noted by John Bargh (1999), much like dandelions, it may be easier to prevent stereotypes from taking root in the first place than to try to pull them up after they are well-established. Interventions aimed at reducing gender essentialism and stereotyping in children are often ineffective (e.g., Gross et al., 2024). However, interventions that target related beliefs or reduce the impact of stereotypes on motivation may be more effective (Master, 2021). Both approaches used in this paper draw on the notion that stereotypes are usually related to essentialist beliefs, the idea that groups have inherent or "essential" properties that cannot change (Ho et al., 2015; Meyer & Gelman, 2016; Mulvey et al., 2010). Psychological essentialism involves the tendency to represent categories in terms of underlying, inherent "essences" that define category membership and are causally responsible for features of that category (Gelman, 2003). Essentialism has been linked to stereotyping among adults (Bastian & Haslam, 2006) and children (Bigler & Liben, 2007). We briefly describe here how each of our two intervention studies was designed to counter stereotypes and then go into further detail in the sections below.

In Study 1, the first approach involved growth mindsets, beliefs that personal characteristics can change (Fig. 1). Teaching students to adopt growth mindsets may reduce the impact of stereotypes by countering the idea that members of each gender have fixed ability in STEM (Dweck & Master, 2009). If characteristics can change, then stereotypes are not set in stone. Anyone could improve their ability in STEM and disprove the stereotype. If girls believe they can grow their abilities in STEM, they may be more motivated to pursue STEM activities, especially in supportive contexts (Burnette et al., 2023; Carroll et al., 2023; Wang et al., 2021).

In Study 2, the second approach, involving structural explanations, rested on children's attributions about whether gender differences are due to internal, stable or external, unstable causes (Fig. 1). For example, if young girls believe that boys are more interested than girls in computer science, with the underlying assumption that this gender difference is inherent and unchanging, then they may reason that they will not be interested in computer science and choose not to pursue it. However, if girls instead believe that gender differences in STEM are highly influenced by social factors and can be shifted (Cheryan et al., 2015, 2017; Valian, 2014), then they may be more open to pursuing computer science and potentially developing an interest in this field. We next discuss both approaches in greater detail.

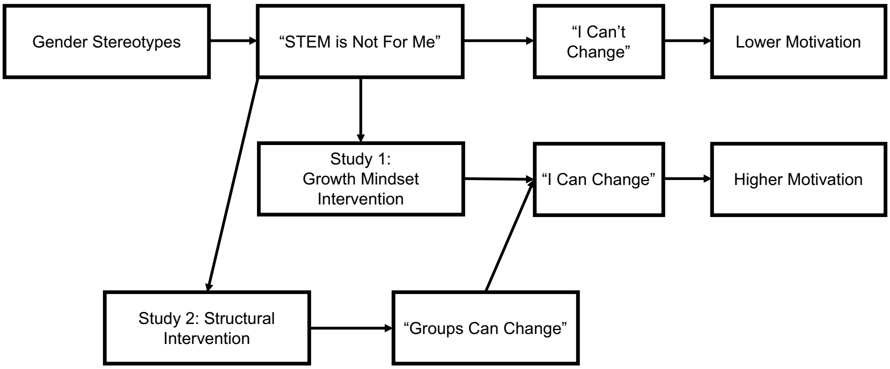

Fig. 1 Theory of Change for Study 1 and Study 2. Gender stereotypes may increase girls' perceptions that STEM is "not for me." In Study 1, girls in the growth mindset condition should be more likely to believe that they could change their ability in computer science, which should lead to higher motivation. In Study 2, girls in the structural differences condition should be more likely than girls in the internal differences condition to agree that groups can change over time because they have an environmental explanation for gender gaps. Agreement that groups can change should lead girls to believe that their personal interest could change, promoting higher motivation

1.3 Growth mindsets as interventions for countering stereotypes

Growth mindsets involve the idea that characteristics are malleable, so groups (and individuals within those groups) have characteristics that are changeable (Dweck & Master, 2009). Previous research on implicit theories or "mindsets" has found that people tend to hold one of two theories that correspond to static versus dynamic conceptions of human nature (Dweck, 2006; Levy, 1999). Those with fixed mindsets (also called "entity theories") believe that people's characteristics cannot change. For example, people who hold a fixed belief about intelligence would agree that people have a certain amount of intelligence and cannot change it. In contrast, other people have growth mindsets (also called "incremental theories") and believe that characteristics can change. For example, people who hold a growth mindset about intelligence would agree that people can always change how intelligent they are. Implicit theories/mindsets can apply to many different characteristics of people, including abilities, interests, personality, and emotions (O'Keefe et al., 2023; Romero et al., 2014).

Growth mindsets about academic ability have important consequences for children's motivation and achievement in school. Students who believe they can improve their own ability through effort are more willing to work hard and take on challenges from which they can learn more (Dweck & Master, 2009). Previous research found that middle-school students who held a growth mindset about intelligence showed higher motivation, including positive beliefs about the value of effort (Blackwell et al., 2007). Middle-school students who were taught the growth mindset showed higher classroom motivation and better math grades compared to students in a control group (see also Yeager et al., 2019). Other findings suggest that growth mindsets correlate with math achievement for children as young as 1st and 2nd grade and predict advanced math course taking in middle school (Park et al., 2016; Romero et al., 2014). Experimentally, high-school students who learned about the struggles of scientists (including their growth and learning over time) changed their stereotypes about scientists and heightened their interest in science compared to learning about scientists' achievements or no background information (Hong & Lin-Siegler, 2012; Lin-Siegler et al., 2016). Even preschool children persisted longer after hearing a story about a scientist who persisted through struggles (Haber et al., 2022).

Because stereotypes involve attributing a fixed set of traits to a group (e.g., "he is a boy so he must like math"), people who hold a fixed mindset should be more likely to endorse stereotypes. In contrast, people who hold a growth mindset should be more likely to acknowledge the dynamic and changing qualities of people (e.g., "he is a boy but his interest in math could change over time"). Indeed, previous research found that adults with a fixed mindset were more likely to endorse stereotypes about ethnic and occupational groups than adults with a growth mindset (Levy et al., 1998). Those with a fixed mindset were also more likely to believe that stereotypes reflected a group's innate character, and less likely to believe that stereotypes arose from environmental or societal influences. The second study in that paper also demonstrated a causal link between mindsets and stereotypes. Specifically, in a randomized experimental study, adults who were taught a fixed mindset about personality were more likely to endorse stereotypes than those taught a growth mindset. People with a fixed mindset are also more likely to maintain stereotypes by paying attention to information that confirms stereotypes and dismissing information that challenges them (Plaks et al., 2001).

Children begin to develop fixed and growth mindsets by age 7 (Gunderson et al., 2013; Heyman et al., 1992; Smiley & Dweck, 1994). Children's mindsets are also related to stereotyping. Children who have fixed mindsets are more likely to: (a) attribute behavior to dispositions rather than the environment, (b) exaggerate differences between groups, (c) minimize differences within groups, and (d) use stereotypes about a group to make a judgment about an unknown member of the group (Levy & Dweck, 1999; Master et al., 2012). Growth mindset interventions have also reduced children's endorsement of gender stereotypes about STEM (Law et al., 2021; see also Lee et al., 2021).

Having a growth mindset about ability can be particularly important for individuals from groups who face negative stereotypes about their ability, such as girls and women in STEM (Dweck & Master, 2009). Stereotypes are more threatening when the negatively stereotyped ability is believed to be fixed (or seen as fixed by others). If a girl believes her STEM ability is fixed, poor performance will confirm the stereotype that girls have less ability in STEM and that she has less ability. This may reduce her sense of self-efficacy and belonging in STEM. Indeed, college women who perceived high levels of gender stereotyping in their math course and felt that the environment was oriented to fixed mindsets reported lower sense of belonging in that class (Good et al., 2012). In contrast, a growth mindset can make stereotypes less threatening, because ability can be improved and stereotypes can be disproved over time (Carr et al., 2012). In one study, a scientist role model who portrayed her ability as having grown over time successfully increased interest and efficacy for 5- to 8-year-old girls of color (Gladstone et al., 2024). Thus, teaching girls to have a growth mindset should lead to improved STEM motivation.

However, these studies represent a limited number of mindset studies that have focused on mindsets about specific STEM abilities, compared to the larger number of studies involving growth mindsets about general intelligence. Some other studies have found that math mindsets are related to math motivation and achievement. One study found that Grade 4 students' fixed math mindsets predicted lower math achievement (Kaya & Karakoc, 2022). Another study found that middle-school students' math mindsets predicted math engagement for students with high metacognitive skills (Wang et al., 2021). However, STEM gender gaps are larger in computer science and engineering compared to math or biology (Cheryan et al., 2017; Master et al., 2025). One study found that a growth mindset intervention among college students increased interest in computer science, but this has not yet been examined with children or adolescents (Burnette et al., 2020). Thus, the focus of Study 1 was whether growth mindsets about computer science might promote girls' motivation for this field in which they are negatively stereotyped.

1.4 Structural explanations for group differences as interventions for countering stereotypes

A second approach to countering stereotypes involves structural explanations, with the goal of promoting attributions that gender differences are due to external and unstable causes, rather than internal and stable causes (Graham & Taylor, 2016). Structural thinking entails explaining a phenomenon by making use of the larger contextual or environmental structure around that phenomenon and identifying contextual constraints that shape the pattern of outcomes within that structure (Vasilyeva & Lombrozo, 2020; Vasilyeva et al., 2018). For example, imagine a town in which children attend either an all-girls school or an all-boys school, and boys are always overrepresented in the town's annual robotics competition (Amemiya & Bian, 2024). A common explanation for this gendered outcome could be that boys are inherently more interested or capable in robotics. However, imagine learning that only the boys' school has a robotics teacher who can help them practice robotics skills. Learning this information should help children attribute the difference in outcomes to a structural difference (the availability of a teacher who could help students practice), rather than an inherent, intrinsic difference.

Thus, considering the larger structure in which choices are made can help to identify a cause of observed differences between social groups—critically, a cause of differences that is malleable (Amemiya et al., 2022; Vasilyeva & Lombrozo, 2020). By changing the structure (e.g., the teachers at the all-girls and all-boys school), we can change the outcomes (e.g., participation in a robotics competition) and reduce differences between groups. In the real world, people frequently make inferences about gendered traits based on the distributions of genders into structural roles, such as occupations (Cejka & Eagly, 1999). Even children are sensitive to cues about inequality and group differences that come from environmental structures (Bigler et al., 2001; Cvencek et al., 2025a; Horwitz et al., 2014; Olson et al., 2011; Zhang et al., 2024).

Thinking about the structural context can be contrasted with internal, essentialist explanations for group differences (Amemiya et al., 2023). For example, one study with adults successfully used the structural context to reduce the effects of stereotype threat (Dar-Nimrod & Heine, 2006). That study involved four conditions. In three conditions, women were told that gender differences in math performance were: (a) due to experiences (locating the source of differences in structures in the external world), (b) due to genetics (locating the source of differences inside each gender in an "essential" way), or (c) non-existent. In the fourth condition, women were primed to think of gender in a standard stereotype threat condition. Women's scores on a math test were significantly higher in the "no differences" and "experiences/structural" compared to the "genetics/essential" and standard stereotype threat conditions. Similarly, another study taught 5- to 8-year-old children about two structural explanations for gender differences—that boys were better than girls at a task because they had more opportunities to practice or that girls were more likely to participate in STEM when they were given more opportunities to do so (Amemiya & Bian, 2024). Children who were provided with one of the two structural explanations were less likely to favor intrinsic explanations for gender differences. If essentialism is a central component of stereotyping, then teaching children to use structural explanations may help to reduce the impact of stereotypes on their motivation (Klysing, 2020).

Stereotypes often rest on beliefs that group differences are due to internal characteristics that cannot change. If children learn that group differences are due to external structures in the environment that can change, then they may discount those stereotypes when evaluating their own motivation. In terms of girls and STEM, if girls learn that gender differences in STEM were caused by factors in the environment, such as different opportunities given to girls and boys, their stereotypes may be weakened. They may be more likely to perceive that such gender differences could change when the environmental factors change as well. If so, they may also reason that their own motivation in STEM could change, if their own environment changed.

1.5 The current paper

The current paper reports findings from two preregistered experimental interventions that sought to counter STEM stereotypes and promote girls' motivation in STEM. In Study 1, middle school students were randomly assigned to watch a video in which they learned about growth mindsets in computer science or a control video involving memory. In Study 2, girls were randomly assigned to watch a video in which they learned that gender differences in STEM interest for a novel activity were due to structural differences in opportunities given to girls and boys or a control video offering inherent explanations.

2 Study 1: growth mindsets

2.1 Rationale

Computer science shows some of the largest disparities among STEM fields (Cheryan et al., 2017). This may be due to the particularly strong gender stereotypes about this field (Master et al., 2021, 2025; Miller et al., 2024). Middle school is a particularly key time to promote girls' motivation in computer science (Code.org et al., 2022; Grover et al., 2016; Rachmatullah et al., 2020). The goal of this study was to examine how a growth mindset intervention affected middle school students' motivation, stereotypes, and performance in computer science. Girls' early experiences in computer science often lead to less positive attitudes compared to boys (Ottenbreit-Leftwich et al., 2021), but growth mindsets may help them remain motivated through difficulties (Dweck & Master, 2009).

2.2 Participants

A total of 381 students from a diverse school district participated in February and March 2022 (219 in Grade 6, 156 in Grade 8, 6 unknown; age range from 11 to 16 years old, *Mage* = 12.71, *SD* = 1.10; 124 girls, 180 boys, 77 non-binary or not reported; 48% Latino/Hispanic, 22% Black/African American, 10% multiracial, 2% Asian, 2% White, 16% other or unknown). As specified in the preregistration, the final analyses excluded participants who failed the attention check and any responses which appeared to be duplicates or non-students. According to the exclusion criteria, 120 participants were excluded from the final analysis, including 116 participants who failed to pass the attention check item, and 14 participants who did not identify their gender as girl or boy. The attention check may have been highly challenging for this sample of participants; in this district in 2022, 78% of students in Grade 6 and 58% of students in Grade 8 did not meet grade level reading standards on statewide standardized tests. The final sample thus consisted of 251 participants (161 in Grade 6, 90 in Grade 8; age range from 11 to 15 years old, *Mage* = 12.67, *SD* = 1.06; 106 girls, 145 boys; 55% Latino/Hispanic; 26% Black/African American, 12% multiracial, 4% Asian, 3% White).

2.2.1 Preregistered power and sample size calculations

Our preregistered target sample size was $N = 252$ (18 students within 14 classrooms). Assuming $\alpha = .05$, power analyses using Optimal Design Plus (Spybrook et al., 2005–2013) for a 2-level linear mixed model with 18 students per classroom (70% consent rate), 14 classrooms (seven per grade level), and a classroom ICC = 0.20, showed that we would have 80% power to detect modest effect sizes ($d = 0.42$ or higher) for student-level predictors and cross-level interactions. For grade level conditional main effects (at Level 2), we would have 80% power to detect large effect sizes ($d = 0.80$ or higher). Final sample size was determined by the number of students who agreed to participate in classes of participating teachers.

2.3 Methods

This study was approved by the University of Washington and University of Houston's Institutional Review Boards and the public school district. Across six middle schools in the southern U.S., teachers were recruited by the school district's technology officer and agreed to participate in the study. All students participated in the district's computer science curriculum. The district curriculum for Grade 6 for Intro to Computer Science involved topics such as sequencing, loops, conditionals, and functions. The district curriculum for Grade 8 for CS Discoveries involved topics such as animation, web design, and AI/machine learning.

Participants completed an online survey in their school during class time while wearing researcher-provided headphones. Participants were randomly assigned through Qualtrics to receive either the ability mindset intervention condition (n = 130) or a memory-related control condition (n = 121) that was equivalent in terms of time, format, and content (see Table 1 and Tables S1 and S2 in the Supplementary Materials). The experimental manipulation involved watching two brief videos telling the story of a middle-school girl named Sophia (introduced by her older sister Mariana) who initially struggled to learn to code (mindset condition) or remember (control condition). The student received support and advice from peers that coding ability can improve (growth mindset intervention) or that memory strategies can be helpful (memory control). The student continued to struggle but remembered the mindset/memory message and was ultimately successful. The memory control condition provided a very conservative test of our hypothesis because it included messages that a characteristic (memory) could change but this cognitive characteristic was not directly relevant to computer science ability. The narration was provided in both audio and visual format simultaneously to maximize comprehension.

Table 1 Study 1 Growth Mindset Experimental Condition

Mindset Condition	Memory-Related Condition
Today, if you want, you can be part of a research project! We are a team of scientists from the University of Houston, and we want to learn more about how kids like you think! We want to learn about your experiences at your school, and how different ways of thinking might help you and your classmates in school. We need your help to make our project better for students in the future	
We are especially interested in what you think about **coding** Coding **is how you put information and commands into a computer, so that you can tell the computer what to do.** You can **make games, apps, programs to help people, or something completely new!**	We are especially interested in what you think about **memory** Memory **means the different ways your brain saves the material you learn in school, so that you can remember it later.** You can **use lots of tricks to help you remember things!**
In this project, you'll watch a short video, and learn about how other students like you feel about coding. Next, we'll ask you some questions about your experiences. There aren't any right or wrong answers. We'll use your ideas to **help future students have a better idea of what coding is like so they can know what to expect**. If you do the activity and questions, you'll be helping us find out more about how kids think and learn	In this project, you'll watch a short video, and learn about what other students like you think about memory. Next, we'll ask you some questions about your experiences. There aren't any right or wrong answers. We'll use your ideas to **help future students have a better idea of how to remember things in coding class so they can know what to do**. If you do the activity and questions, you'll be helping us find out more about how kids think and learn
Conversation between Sophia and her sister Mariana emphasizes: **Learning more** about coding **and trying** can help you **understand** coding **better and get better at it**. When students start using that **growth** mindset, they can **learn to be better at** coding	Conversation between Sophia and her sister Mariana emphasizes: **Using different memory strategies** can help you **remember more and develop a better memory**. When students start using that **memory** mindset, they can **use tricks to remember**
If you were Sophia's classmate, what would you say to her?	
Tell her about a time in your life when you thought you were not **good at** something in school but became **better at** it later. Try to choose an example where you really weren't **good at** it at first but became better **at it** over time	Tell her about a time in your life when you thought you **couldn't remember** something in school but **figured out how to remember** it. Try to choose an example where you really **couldn't remember** at first but **started to remember** over time
Conversation between Sophia, Mariana, and peers emphasizes: Using **growth mindsets,** they got better at coding over time	Conversation between Sophia, Mariana, and peers emphasizes: Using **memory strategies,** they got better at remembering over time
How is it helpful for students to know that **abilities** can **get better** over time? Is there anything you want future students at your school to know about how to become **better at** coding?	How is it helpful for students to think with this mindset, that it's important to **use different memory strategies**? Is there anything you want future students at your school to know about **memory for** coding?

Note. Bold text indicates language that differed by condition

2.3.1 Measures

The survey was administered at the end of the experimental manipulation video. The survey included: (a) an attention check requesting that participants mark a particular response, (b) questions about responding to the videos, (c) ability and interest mindsets; (d) belief in gender-interest and gender-ability stereotypes; (e) four motivational variables: interest, sense of belonging, ability self-concept, and identification for computer science classes and activities in school; (f) performance on a coding assessment; and (g) demographics (gender, race/ethnicity, and birth date). The survey also included other measures (e.g., open-ended questions), but the present study focused on hypotheses involving quantitative analyses. The attention check came immediately after the mindset measures and instructed students, "*Please choose the circle marked 'slightly disagree' to show that you read this question.*"

Mindsets included both ability mindsets and interest mindsets, measured by three items each, using a 6-point Likert scale from 1 (*Strongly disagree*) to 6 (*Strongly agree*), with the order of mindset scales counterbalanced. Ability mindsets measured how much participants believed their coding abilities can be changed (e.g., *"Being good at coding is something that you can't change very much"*) and interest mindsets measured how participants believed their interest in coding can be changed (e.g., *"People can't change whether they like coding or not"*). Mindsets were coded such that higher values represented more fixed mindsets. The ability mindset scales showed low internal reliability, $\alpha = .62$. Interest mindsets showed low internal reliability, $\alpha = .54$, and were measured by three items, including a newly-developed item about belief about whether high potential interest is universal or limited to only some children (*"Only some kids have the potential to become really interested in coding later on in life;"* Rattan et al., 2012). However, that item reduced the overall reliability of the scale. Therefore, we dropped that item and used the 2-item interest mindset scale for the final analyses, $\alpha = .61$. Cronbach's alpha is often lower among child participants and on scales with fewer items (Taber, 2018). Because these reliabilities were fair but not good for short-length scales (Xiao et al., 2024), we examined the inter-item correlations as well. The mean inter-item correlation for the ability mindset scale was $r = .36$ (range: .26–.44), and the inter-item correlation for the interest mindset scale was $r = .44$, falling within the recommended range of .15–.50 for demonstrating internal consistency (Clark & Watson, 1995).

Four motivation variables, including interest, sense of belonging, ability self-concept, and identification, were measured by 6-point Likert scales from 1 (*Strongly disagree*) to 6 (*Strongly agree*). Interest assessed participants' interest in computer coding activities with three items (e.g., *"I like to do coding activities"*), showing good reliability, $\alpha = .94$. Sense of belonging assessed participants' perceptions of belonging in computer coding class with three items (e.g., *"I feel like I belong in my coding class"*), showing good reliability, $\alpha = .81$. Ability self-concept assessed participants' beliefs about their computer coding abilities with three items (e.g., *"I am good at coding activities"*), showing good reliability, $\alpha = .88$. Coding identification assessed how strongly participants identified themselves with computer science with six items (e.g., *"I am a coding person"*), with good reliability, $\alpha = .91$.

Gender stereotypes about computer science were assessed by measuring participants' beliefs about boys' and girls' interests (e.g., *"How much do you think that most [boys/girls] like coding?"*) and ability (e.g., *"How good do you think most [boys/girls] are at coding?"*) on Likert scales from 1 (*Really do not like/not good*) to 6 (*Really like/good*), respectively. Each type of gender stereotype about computer science was then computed as difference scores, with ratings of boys – ratings of girls (Master et al., 2021). Positive scores indicated stereotypical beliefs that boys were more interested in or better at computer science; conversely, negative scores indicated beliefs that girls were more interested in or better at computer science.

Participants' coding performance was assessed by measuring their coding knowledge on a 10-item scale calibrated to the Grade 6 curriculum (e.g., *"There is one extra block that is going to cause the bird to crash. Which block needs to be unhooked for the program to run successfully?"*) or the Grade 8 curriculum (e.g., *"Why is it important for a programmer to write out an algorithm before trying to program it?"*). The computer coding assessment was developed by the school district technology director based on the Code.org curricula for CS Fundamentals (Grade 6) and CS Discoveries (Grade 8). The final score was computed as the sum of correct responses to the ten questions (range 0 to 10).

2.3.2 Preregistration

The hypotheses and analyses plan were preregistered on the Open Science Framework at https://osf.io/pr8fc/files/y4xfuhttps://osf.io/pr8fc/?view_only=6c82ef9621d4422eb27a0484f5faf155. Materials are available at https://osf.io/peu3x/overviewhttps://osf.io/peu3x/overview?view_only=06a7554511d94b7e84494ba6cd677993.

This preregistration indicated a plan to use linear mixed models by classroom. Because we did not receive data about specific classrooms from the district, we tested models with students nested within teachers. However, these models showed multiple negative ICCs (for ability mindsets, interest stereotypes, and ability stereotypes), which likely indicated that the between-group variance was essentially near zero. Because the ICC represents the proportion of total variance attributable to between-group differences, it cannot be negative in the population. Negative empirical estimates are therefore generally taken to reflect a between-group variance of effectively zero (or model misspecification) and are conventionally treated as zero (Snijders & Bosker, 2011; Wang et al., 2011). The dataset included only six schools and nine teachers, so it is likely that the random effects estimation was unstable and underpowered for a multilevel structure. With so few level-2 units, maximum-likelihood variance components tend to be downwardly biased and imprecise, and the sampling distribution of a near-zero variance component is highly skewed (Heck & Thomas, 2020). When within-group observations are effectively independent, the ICC approaches zero and the multilevel model reduces to a single-level specification (Wang et al., 2011). Below, we report the single-level models for ability mindsets, interest stereotypes, and ability stereotypes; we report results from multilevel models for all other outcome variables.

Table 2 Study 1 Correlations by Condition

Variable	1	2	3	4	5	6	7	8	9	10
1. Ability Mindset		.39**	−.21*	−.18*	−.21*	−.13	.04	.14	.02	−.39*
2. Interest Mindset	.44**		−.13	−.14	−.11	−.10	.15	−.004	−.07	−.34*
3. Interest	−.14	−.05		.74**	.74**	.77**	.02	.05	.12	.32*
4. Belonging	−.18**	−.10	.80**		.78**	.75**	−.02	.07	.17	.31*
5. Self-concept	−.15	.02	.74**	.71**		.71**	.005	.02	.14	.31*
6. Identification	−.13	−.05	.78**	.73**	.69**		.004	.03	.06	.23
7. Ability stereotype diff. score	−.02	.04	.02	.09	.03	.07		.29**	−.11	.24
8. Interest stereotype diff. score	−.05	−.01	.14	.17	.05	.17	.57**		.07	.26
9. Performance (Grade 6)	−.11	−.18	−.17	−.07	.06	−.07	−.05	−.08		–
10. Performance (Grade 8)	−.05	.01	.37**	.38*	.37*	.42**	−.003	.23	–	

Note. Correlations for variables in the ability mindset condition are displayed below the diagonal (N = 130). Correlations for variables in the control condition are displayed above the diagonal (N = 121). Diff. score = difference score (ratings about boys minus ratings about girls)

$^{*}p < .05$, $^{**}p < .01$

Table 3 Study 1 Correlations by Grade Level

Variable	1	2	3	4	5	6	7	8	9
1. Ability Mindset		.47**	−.37**	−.36**	−.33**	−.31**	−.14	−.12	−.25*
2. Interest Mindset	.38**		−.14	−.12	−.01	−.04	.03	.07	−.17
3. Interest	−.03	−.07		.81**	.80**	.77**	.17	.26**	.34**
4. Belonging	−.05	−.14	.68**		.77**	.76**	.23*	.34**	.35**
5. Self-concept	−.09	−.09	.59**	.65**		.76**	.23*	.33*	.34**
6. Identification	−.02	−.10	.76**	.69**	.59**		.20	.27*	.34**
7. Ability stereotype diff. score	.11	.14	−.10	−.13	−.16*	−.08		.66**	.12
8. Interest stereotype diff. score	−.002	−.05	.02	−.004	−.15	.02	.23**		.23*
9. Performance	−.06	−.14	−.04	.04	.09	−.02	−.08	−.003	

Note. Correlations for variables for students in Grade 6 are displayed below the diagonal (N = 161). Correlations for variables for students in Grade 8 are displayed above the diagonal (N = 90). Diff. score = difference score (ratings about boys minus ratings about girls)

$^{*}p < .05$, $^{**}p < .01$

2.4 Results

2.4.1 Correlations

For correlations between all variables split by condition (mindset condition or control), grade (Grade 6 or 8), and gender (girls or boys), see Tables 2, 3, 4. Fixed mindsets about ability and interest were significantly correlated for students in both conditions, both grade levels, and both girls and boys, ps < .01. Students in both conditions who reported more fixed ability mindsets also reported a lower sense of belonging in coding, ps < .05. Students in Grade 8 with more fixed ability mindsets also reported lower interest in coding, a lower sense of belonging, lower ability self-concepts, and lower identification with coding, ps < .01. They also had lower performance on the coding assessment, p < .05. Boys with more fixed ability mindsets also reported lower interest, lower sense of belonging, lower self-concepts, and lower

Table 4 Study 1 Correlations by Gender

Variable	1	2	3	4	5	6	7	8	9	10
1. Ability Mindset		.48**	−.06	−.07	−.08	−.02	.09	.13	.07	.008
2. Interest Mindset	.36**		.12	.03	.12	.13	.10	.14	.07	−.03
3. Interest	−.22**	−.23**		.84**	.80**	.81**	.14	.05	−.06	.11
4. Belonging	−.22**	−.19*	.69**		.83**	.77**	.17	.06	−.12	−.05
5. Self-concept	−.21**	−.12	.64**	.65**		.71**	.15	.07	−.04	.05
6. Identification	−.18**	−.18*	.74**	.71**	.66**		.08	−.05	.03	.02
7. Ability stereotype diff. score	−.002	.10	−.10	−.07	−.11	−.02		.64**	.05	.23
8. Interest stereotype diff. score	−.14	−.08	.09	.13	−.03	.16	.34**		.12	.23
9. Performance (Grade 6)	−.14	−.26*	−.10	.13	.14	−.10	−.13	−.06		–
10. Performance (Grade 8)	−.29*	−.24	.37**	.43**	.40**	.42**	.005	.17	–	

Note. Correlations for boys are displayed below the diagonal (N = 145). Correlations for girls are displayed above the diagonal (N = 106). Diff. score = difference score (ratings about boys minus ratings about girls)

$^{*}p < .05$, $^{**}p < .01$

identification, ps < .05. Boys in Grade 6 with more fixed interest mindsets scored lower on the coding assessment, $p < .05$, while boys in Grade 8 with a higher ability fixed mindset scored lower on the coding assessment, $p < .05$.

2.4.2 Descriptive statistics

Descriptive statistics for all variables by condition and grade level can be found in Table 5.

2.4.3 Mindsets

We analyzed ability mindsets, gender-interest stereotypes, and gender-ability stereotypes with ANOVA because their intraclass correlations were near zero (see Table S3), and we analyzed the remaining variables with multilevel models in which students were nested within teachers (see Table S4). For the ANOVA, we conducted a 2 (condition: ability mindset or control) × 2 (grade level: 6 or 8) × 2 (gender: girl or boy) model. For the multilevel models, we included the main effects of condition, grade level, and gender, the condition by gender and grade by gender interactions, and the three-way interaction. All condition by grade interactions were nonsignificant and removed to simplify the multilevel models.

First, for ability mindsets, the interaction between condition, gender, and grade level was not significant, $F(1, 243) = 0.54$, $p = .47$, $\eta_p^2 = 0.002$. There were no interaction effects between condition and grade, $F(1, 243) = 2.40$, $p = .12$, $\eta_p^2 = 0.01$, between condition and gender, $F(1, 243) = 0.38$, $p = .54$, $\eta_p^2 = 0.002$, or between grade and gender, $F(1, 243) = 2.82$, $p = .095$, $\eta_p^2 = 0.01$. There was a significant main effect of gender, $F(1, 243) = 7.91$, $p = .005$, $\eta_p^2 = 0.03$. Girls reported more fixed ability mindsets than boys. There was no significant main effect of condition, $F(1, 243) = 1.44$, $p = .23$, $\eta_p^2 = 0.006$, or grade level, $F(1, 243) = 0.30$, $p = .59$, $\eta_p^2 = 0.001$. Examining the simple effects of condition within gender, there was no significant

Table 5 Descriptive Statistics for Study 1

Grade	Variable	Ability Mindset Condition			Control Condition		
		N	*M*	*SD*	*N*	*M*	*SD*
Grade 6	1. Ability Mindset						
	Boys	48	2.47	1.03	41	2.43	1.11
	Girls	39	2.65	1.18	35	2.58	0.99
	2. Interest Mindset						
	Boys	48	2.69	1.11	41	2.33	1.18
	Girls	39	2.72	1.16	35	2.84	0.88
	3. Interest						
	Boys	48	5.08	0.79	41	5.11	0.67
	Girls	39	4.89	0.81	35	4.41	1.05
	4. Sense of Belonging						
	Boys	48	4.51	0.81	41	4.56	0.91
	Girls	39	4.40	0.81	35	3.84	1.03
	5. Self-concept						
	Boys	48	4.97	0.87	41	5.07	0.74
	Girls	39	4.64	0.76	35	4.25	1.10
	6. Identification						
	Boys	48	4.14	1.09	41	4.26	0.94
	Girls	39	3.93	1.05	35	3.54	1.08
	7. Ability stereotypes diff. score						
	Boys	48	0.10	0.99	41	0.07	1.06
	Girls	39	-0.03	0.49	35	0.09	0.45
	8. Interest stereotypes diff. score						
	Boys	48	0.13	1.06	41	0.00	1.43
	Girls	39	0.10	0.72	35	0.26	1.01
	9. Performance						
	Boys	48	5.85	1.87	41	6.12	1.71
	Girls	39	5.51	1.80	35	5.63	1.35

Table 5 (continued)

Grade	Variable	Ability Mindset Condition			Control Condition		
		N	*M*	*SD*	*N*	*M*	*SD*
Grade 8	1. Ability Mindset						
	Boys	30	2.18	1.05	26	2.38	1.10
	Girls	13	2.64	1.24	19	3.25	0.92
	2. Interest Mindset						
	Boys	30	2.58	1.16	26	2.54	1.01
	Girls	13	2.38	1.19	19	2.71	1.00
	3. Interest						
	Boys	30	4.46	1.37	26	4.65	1.07
	Girls	13	3.59	1.53	19	3.32	1.35
	4. Sense of Belonging						
	Boys	30	4.04	1.33	26	3.97	1.40
	Girls	13	2.95	1.34	19	3.11	1.13
	5. Self-concept						
	Boys	30	4.33	1.21	26	4.33	1.20
	Girls	13	3.03	1.58	19	3.23	1.30
	6. Identification						
	Boys	30	3.72	1.39	26	3.40	1.37
	Girls	13	3.12	1.43	19	2.68	1.26
	7. Ability stereotypes diff. score						
	Boys	30	0.40	1.10	26	0.31	1.05
	Girls	13	-0.31	0.85	19	0.00	1.05
	8. Interest stereotypes diff. score						
	Boys	30	0.60	1.13	26	0.69	1.29
	Girls	13	-0.15	1.28	19	0.21	1.18
	9. Performance						
	Boys	30	6.26	1.81	26	6.29	1.85
	Girls	13	6.23	1.44	19	4.80	1.36

effect of condition for girls, $F(1, 243) = 1.35$, $p = .25$, $\eta_p^2 = 0.006$, or for boys, $F(1, 243) = 0.22$, $p = .64$, $\eta_p^2 = 0.001$.

Second, for interest mindsets, we fit a multilevel model with students nested within teachers (see Table S4). The condition by gender interaction was not significant, $b = 0.23$, $SE = 0.31$, $p = .45$, nor was the three-way interaction, $b = 0.38$, $SE = 0.46$, $p = .40$. There were no significant main effects of condition, $b = -0.21$, $SE = 0.18$, $p = .23$, grade level, $b = 0.02$, $SE = 0.18$, $p = .93$, or gender, $b = 0.11$, $SE = 0.22$, $p = .61$. The simple effects of condition within gender were not significant for girls, $b = -0.21$, $SE = 0.23$, $t = -0.91$, $p = .36$, or for boys, $b = 0.21$, $SE = 0.18$, $t = 1.21$, $p = .23$.

2.4.4 Motivation

Motivation was examined in terms of condition, gender, and grade level differences for each measure.

For interest, we fit a multilevel model with students nested within teachers. The condition by gender interaction was significant, $b = -0.59$, $SE = 0.29$, $p = .037$. Examining the simple effects of condition within gender, girls in the mindset condition reported higher interest in coding activities than girls in the control condition, $b = 0.43$, $SE = 0.21$, $t = 1.99$, $p = .047$, whereas there was no effect of condition for boys, $b = -0.11$, $SE = 0.17$, $t = -0.68$, $p = .50$. The grade by gender interaction was marginally significant, $b = -0.66$, $SE = 0.36$, $p = .068$. Girls in Grade 6 reported higher interest than girls in Grade 8, $b = 1.02$, $SE = 0.24$, $t = 4.19$, $p < .001$, and boys in Grade 6 reported higher interest than boys in Grade 8, $b = 0.40$, $SE = 0.20$, $t = 2.05$, $p = .04$. Boys reported higher interest than girls in both Grade 6, $b = 0.48$, $SE = 0.16$, $t = 3.04$, $p = .002$, and Grade 8, $b = 1.10$, $SE = 0.22$, $t = 4.92$, $p < .001$. The three-way interaction was not significant, $b = 0.09$, $SE = 0.43$, $p = .84$.

For sense of belonging, the condition by gender interaction was significant, $b = -0.57$, $SE = 0.29$, $p = .047$. Girls in the mindset condition reported a higher sense of belonging than girls in the control condition, although this simple effect was not significant, $b = 0.24$, $SE = 0.22$, $t = 1.13$, $p = .26$, and there was no effect of condition for boys, $b = -0.01$, $SE = 0.17$, $t = -0.07$, $p = .95$. The grade by gender interaction was significant, $b = -0.80$, $SE = 0.37$, $p = .03$. Girls in Grade 6 reported a higher sense of belonging than girls in Grade 8, $b = 0.77$, $SE = 0.26$, $t = 3.00$, $p = .003$, whereas the difference between boys in Grade 6 and Grade 8 was not significant, $b = 0.28$, $SE = 0.20$, $t = 1.38$, $p = .17$. Boys reported a higher sense of belonging than girls in both Grade 6, $b = 0.47$, $SE = 0.16$, $t = 2.96$, $p = .003$, and Grade 8, $b = 0.96$, $SE = 0.23$, $t = 4.24$, $p < .001$. The three-way interaction was not significant, $b = 0.62$, $SE = 0.44$, $p = .16$.

For ability self-concepts, the condition by gender interaction was marginally significant, $b = -0.47$, $SE = 0.28$, $p = .090$, and the simple effects of condition within gender were not significant for girls, $b = 0.19$, $SE = 0.21$, $t = 0.91$, $p = .37$, or for boys, $b = -0.09$, $SE = 0.16$, $t = -0.54$, $p = .59$. The grade by gender interaction was marginally significant, $b = -0.69$, $SE = 0.36$, $p = .055$. Girls in Grade 6 reported higher ability self-concepts than girls in Grade 8, $b = 0.83$, $SE = 0.25$, $t = 3.31$, $p = .001$, whereas the difference between boys in Grade 6 and Grade 8 was not significant, $b = 0.32$, $SE = 0.20$, $t = 1.61$, $p = .11$. Boys reported higher ability self-concepts than girls in both Grade 6, $b = 0.61$, $SE = 0.16$, $t = 3.92$, $p < .001$, and Grade 8, $b = 1.12$, $SE = 0.22$, $t = 5.03$, $p < .001$. The three-way interaction was not significant, $b = 0.36$, $SE = 0.43$, $p = .40$.

For identification, the condition by gender interaction was not significant, $b = -0.36$, $SE = 0.32$, $p = .26$. However, the simple effect of condition within girls was significant, with girls in the mindset condition identifying more strongly with coding than girls in the control condition, $b = 0.49$, $SE = 0.24$, $t = 2.04$, $p = .04$, whereas there was no effect of condition for boys, $b = -0.02$, $SE = 0.19$, $t = -0.11$, $p = .91$. The grade by gender interaction was not significant, $b = -0.01$, $SE = 0.41$, $t = -0.11$, $p = .99$. Boys identified more strongly as a coding person than girls in both Grade 6, $b = 0.49$, $SE = 0.18$, $t = 2.79$, $p = .005$, and Grade 8, $b = 0.63$, $SE = 0.25$, $t = 2.48$, $p = .013$. The three-way interaction was not significant, $b = -0.25$, $SE = 0.48$, $p = .60$.

2.4.5 Stereotypes

We used a 2 (condition: ability mindset or control) × 2 (grade level: Grade 6 or 8) × 2 (gender, girls or boys) ANOVA to examine stereotype difference scores.

For gender-interest stereotypes, the interaction between condition, grade level, and gender was not significant, $F(1, 243) = 0.00$, $p = .99$, $\eta_p^2 = 0.00$. We found a significant interaction effect between grade level and gender, $F(1, 243) = 5.58$, $p = .02$, $\eta_p^2 = 0.02$. Boys in Grade 8 were more likely than boys in Grade 6 to report stereotypes favoring boys, $F(1, 243) = 9.04$, $p = .003$, $\eta_p^2 = 0.04$. Boys in Grade 8 were more likely than girls in Grade 8 to report stereotypes favoring boys, $F(1, 243) = 5.89$, $p = .016$, $\eta_p^2 = 0.02$. The interactions between condition and grade level, $F(1, 243) = 0.47$, $p = .49$, $\eta_p^2 = 0.002$, and between condition and gender, $F(1, 243) = 0.79$, $p = .38$, $\eta_p^2 = 0.003$, were not significant. We found no significant main effect based on condition, $F(1, 243) = 0.61$, $p = .44$, $\eta_p^2 = 0.003$. Examining the simple effects of condition within gender, there was no significant effect of condition for girls, $F(1, 243) = 1.14$, $p = .29$, $\eta_p^2 = 0.005$, or for boys, $F(1, 243) = 0.01$, $p = .93$, $\eta_p^2 = 0.00$.

For gender-ability stereotypes, the interaction between condition, grade level, and gender was not significant, $F(1, 243) = 0.27$, $p = .60$, $\eta_p^2 = 0.001$. There was a significant main effect of gender, $F(1, 243) = 5.22$, $p = .023$, $\eta_p^2 = 0.021$. Boys were more likely than girls to report stereotypes favoring boys. The interaction between grade level and gender was marginally significant, $F(1, 243) = 3.28$, $p = .07$, $\eta_p^2 = 0.01$. The main effect of condition was not significant, $F(1, 243) = 0.36$, $p = .55$, $\eta_p^2 = 0.001$. The interactions between condition and grade level, $F(1, 243) = 0.07$, $p = .79$, $\eta_p^2 = 0.00$, and between condition and gender, $F(1, 243) = 1.20$, $p = .28$, $\eta_p^2 = 0.005$, were not significant. Examining the simple effects of condition within gender, there was no significant effect of condition for girls, $F(1, 243) = 1.17$, $p = .28$, $\eta_p^2 = 0.005$, or for boys, $F(1, 243) = 0.16$, $p = .69$, $\eta_p^2 = 0.001$.

2.4.6 Performance

For coding performance, we fit multilevel models with students nested within teachers separately for the Grade 6 and Grade 8 assessments, including condition, gender, and the interaction between condition and gender. For the Grade 6 assessment, the condition by gender interaction was not significant, $b = -0.11$, $SE = 0.51$, $p = .83$, and there were no significant effects of condition, $b = 0.27$, $SE = 0.34$, $p = .44$, or gender, $b = -0.26$, $SE = 0.35$, $p = .46$. Examining the simple effects of condition within gender, girls in the ability growth mindset condition did not differ from girls in the control group, $b = -0.16$, $SE = 0.37$, $t = -0.42$, $p = .67$, and boys in the mindset condition did not differ from boys in the control group, $b = -0.27$, $SE = 0.34$, $t = -0.77$, $p = .44$.

For the Grade 8 assessment, the condition by gender interaction was marginally significant, $b = -1.32$, $SE = 0.70$, $p = .059$. Examining the simple effects of condition within gender, girls in the ability growth mindset condition scored higher than girls in the control group, $b = 1.37$, $SE = 0.55$, $t = 2.50$, $p = .013$, whereas there was no effect of condition for boys, $b = 0.06$, $SE = 0.41$, $t = 0.14$, $p = .89$. In the control group, boys scored higher than girls, $b = 1.24$, $SE = 0.47$, $t = 2.62$, $p = .009$.

2.5 Discussion

Study 1 examined the effects of a growth mindset manipulation on middle school students' motivation, stereotypes, and performance in computer science. In multilevel models, the condition by gender interaction was significant for interest and sense of belonging and marginally significant for ability self-concepts. Notably, the simple effects of condition for girls were significant for interest and identification, with girls in the mindset intervention condition reporting higher motivation than girls in the control condition, and Grade 8 girls in the mindset intervention condition scored higher on a coding test than girls in the control condition. In addition, correlational results supported the general importance of promoting growth mindsets and positive motivational beliefs during this period. Among students in the control condition, fixed ability mindsets were linked to lower interest, sense of belonging, ability self-concepts, and performance in computer science.

Using the mindset measure as a manipulation check of whether the experimental manipulation caused differences in participants' mindsets, the intervention did not successfully increase students' growth mindsets about computer science. Why was our intervention ineffective in promoting growth mindsets? We offer five possibilities each with implications for future research. One possibility is that the narrative video format weakened the intervention compared to other growth mindset interventions that involved reading information that emphasized scientific evidence for the idea of growth mindsets (e.g., Yeager et al., 2019). The video format was intended to make the manipulation more engaging and relatable for students, but students may have found the information less credible or trustworthy, or the animation too childish. A second possibility is that the control condition created an extremely conservative test of the growth mindset effect, because it also involved a cognitive ability that could be improved (memory). Although previous growth mindset manipulations showed improvement compared to memory-related control conditions, our video format may have led to similar perceptions of the videos in the two conditions. In support of this option, the means for mindsets in both conditions indicated slight to medium support of growth mindsets (values between 2 and 3 on the scale, where 1 indicated strong growth mindsets and 6 indicated strong fixed mindsets).

A third possibility involves the domain-specific vs. domain-general nature of mindsets. The current study used a domain-specific version of growth mindsets (that ability in computer science can improve). This was intended to promote domain-specific motivation. However, students' general mindsets about the malleability of intelligence may continue to play an important role in their motivation across domains, making domain-specific effects more difficult to capture. A fourth possibility involves the timing of the manipulation. Students were well into the semester and had many weeks experience with coding by the time of the study. These prior experiences with computer science may have already solidified their perceptions of the malleability of computer science skills. Their classroom climates may also have contributed to prior experiences that impacted the effectiveness of the intervention. Given students' general endorsement of growth mindsets in the control condition, teachers and peers may already have given signals indicating that coding ability was malleable (Yeager et al., 2019). One brief animated video may have failed to make a measurable impact against the weight of these past real-world experiences. Finally, it is also important to note that the mindset measures showed low reliability ($<.70$), which impacted the effectiveness of this measure. Thus, future studies should aim to develop more reliable measures of students' mindsets about computer science with this age group.

Results also showed consistent interaction patterns between gender and grade level. Gender gaps widened between Grade 6 and Grade 8, with Grade 8 girls showing lower motivation for computer science compared to Grade 6 girls for interest, sense of belonging, and ability self-concepts. This pattern indicates that middle school is a key time to focus on promoting girls' motivation for computer science, particularly providing supportive environments in foundational computer science classes. Although there may be multiple underlying sources of this drop in motivation for girls, future research should examine the classroom climate within middle school classrooms.

In terms of the proposed theory of change (Fig. 1), different types of motivation may be more or less likely to shift following beliefs about change, with stable trait-like motivation measures being less likely to shift compared to more contextually-specific measures. The motivation items generally assessed motivation as a stable trait (e.g., "*I like to do coding activities*," and "*I am a coding person*"). It is possible that such measures are less sensitive to change based on mindset beliefs. Future research should consider examining whether other types of motivational measures may be more malleable and sensitive to this type of intervention. For example, task-specific self-efficacy may be more likely to shift following a growth mindset intervention that is also more closely linked to that particular task (e.g., using physical devices in computing with the micro:bit device; Gál & Kerekes, 2025). In other words, students' beliefs that they could improve their ability to use physical computing devices may have more direct effects on their self-efficacy for using those devices. Similarly, future studies could consider measuring behavioral, cognitive, emotional, and agentic engagement during a particular learning activity as more situationally sensitive measures (Reeve et al., 2025).

The results from the stereotype measures showed that boys were more likely than girls to endorse gender stereotypes favoring boys in computer science. These stereotypes may have contributed to a "chilly" classroom climate for girls through the comments and behavior of boys, leading to a stronger sense of social identity threat or stereotype threat (Walton et al., 2015). The social context may play a large role in the effectiveness of motivational interventions (Hecht et al., 2023). For example, peer norms that support challenge seeking can increase the effectiveness of growth mindset interventions (Yeager et al., 2019). In terms of girls and coding, there may be social costs for girls if peers find their interest in coding surprising or "unfeminine," making interest harder to shift (similar to costs that have been shown for undergraduate women; Cheryan et al., 2020). Thus, future research could also examine whether growth mindset interventions show more long-term positive effects in combination with interventions designed to reduce stereotyping among boys (e.g., Cyr et al., 2024). They could also track the effectiveness of mindset interventions in the context of peer and teacher mindsets (Hecht et al., 2022; Yeager et al., 2019, 2022).

The effectiveness of the intervention may also have been affected by participants' other social identities, particularly the intersectionality between gender and race/ethnicity. Students with multiple marginalized identities (e.g., Black and Latina girls) can experience interlocking systems of oppression and bias that can differ from the experiences of White girls (Rodriguez & Blaney, 2021). The video protagonist was designed to seem Latina, which matched the most common demographic group in the school district (55% of our final sample), and her friends in the video looked like they came from diverse racial/ethnic backgrounds. She was designed to serve as an effective role model who could normalize struggle and improvement (Barr et al., 2024; Herrmann et al., 2016; Shin et al., 2016). Research indicates that role models help to counter gender stereotypes and support girls' motivation in coding (Cheryan et al., 2013a, b; Stout et al., 2011). However, research also suggests that similarity to role models is important, including racial/ethnic similarity (Van Camp et al., 2019). The intervention may have been more effective for Latina girls than for Black girls, although we did not have the statistical power to test for race/ethnicity as an additional moderator.

Despite the lack of an overall effect on the mindset manipulation check, there was some evidence that the intervention benefited girls. The simple effects for girls showed increases in coding interest and identification. There was also a significant effect of condition on the performance of girls in Grade 8 on a computer science assessment. Due to the small cell size (n = 32) and the failure of the manipulation check, these findings should be replicated and should not be interpreted as caused by the growth mindset manipulation. For example, it is possible that the intervention did not affect girls' own mindsets, but their perceptions that others had growth mindsets (Hecht et al., 2022; Yeager et al., 2022).

3 Study 2: structural explanations

3.1 Rationale

Stereotypes reinforce beliefs that gender differences are due to internal and inherent gender differences, minimizing recognition of environmental influences. What kinds of messages might help disrupt this process? If girls learn how environmental structures have contributed to gender gaps in interest, and that those structures can change, this knowledge may reduce the impact of stereotypes on their own interest in computer science. We examined consequences of teaching girls about external, rather than internal, causes of gender gaps (see Table 6). Our objective was to examine whether giving 11- to 12-year-old girls structural explanations (as one specific type of external explanation) for gender differences affected their stereotypical beliefs and interest in a novel STEM activity. A novel STEM activity was chosen to avoid participants' preexisting beliefs about familiar STEM domains (Yang et al., 2022).

3.2 Participants

Participants were 220 11- and 12-year-old girls (71% White, 19% multiracial, 6% Asian, 4% Other; *Mage* =12.02, *SD* = 0.59) who participated between November 2020 and June 2021. The research included only children who self-identified as girls. They were recruited from families who had signed up for a research subject list or online participant pools. We preregistered that participants would be excluded and replaced if they were extremely distracted by their home surroundings (rated as "very distracted" by the researcher), showed obvious fatigue, had difficulty answering the

Table 6 Study 2 Structural Condition Experimental Manipulation

	Control Condition	Structural Condition
	Today we're going to talk about a new activity called the finding game. In the finding game, you just find the block that matches the block at the top. The finding game looks like this. If you rotate the block at the top, and tilt it forward, it looks just like the block on the right. The other two blocks are both a little bit different. Do you see how it works? I want to tell you something important about the finding game. Boys are more likely to play the finding game than girls are. Let's talk more about WHY boys are more likely to play the finding game than girls are.	
	It's not because of the outside world. It's because of what boys and girls like to do. People like to do different things. Some kids like to play the finding game and some kids like to play other games more. Different kids like different things.	It's not because boys and girls are naturally different. It's because of how things are set up in the world. Adults treat girls and boys differently. Boys are given more opportunities to play the finding game, and girls are given more other games instead. Different experiences made them different.
	Lots of toy stores have separate aisles for girls' and boys' toys. Did you know that the section with boys' toys is more likely to have games like the finding game? Now imagine a boy and a girl walk into the store. They each buy a toy. Who do you think is more likely to buy the finding game?	
	In real life, this happens over and over. Boys like to play lots more with games like the finding game than girls do. That's why boys say they like the finding game more than girls do. It's not because of the outside world. It's because kids are different. Different kids like different things. That's just the way it is. Sometimes people in the world just like different things. They do different activities because they like different things. Even if you change things in the outside world, the kids are still different.	In real life, this happens over and over. Boys get lots more chances to play with games like the finding game than girls do. That's why boys say they like the finding game more than girls do. It's not because girls and boys were born different. It's because they were treated differently. Different experiences made them different. BUT THAT CAN CHANGE. Sometimes things in the world are set up in ways that are not fair for different groups. Those outside differences can CAUSE differences between groups, when there were no differences before. But if you take away those outside things, the differences between groups go away.

questions (2 additional participants excluded), completed the study on a smartphone rather than a larger device (3 additional participants excluded), had technical difficulties and were unable to see/hear parts of the study (2 additional participants excluded), or reported suspicion about or previous familiarity with the study. An additional 12 participants were excluded because they self-identified their gender as non-binary rather than as girls, for a final sample of $N = 220$.

3.2.1 Preregistered power and sample size calculations

We used G*power 3.1 to calculate the sample size to detect a simple effect of condition on interest, using the effect size $d_s = .40$ from a pilot study. Our power calculation suggested obtaining a sample size of $N = 216$ girls to detect this effect. To account for potential overestimates of effect size, we planned to stop data collection at 220 participants. We then used G*Power 3.1 to calculate the sample size to detect a simple effect of condition on number of activities chosen, using the effect size $d_s = .80$ from a pilot study. Our power calculation suggested obtaining a sample size of $N = 56$ girls to detect this effect.

3.3 Methods

3.3.1 Experimental manipulation

Participants completed the activity online with a researcher using Zoom. The researcher shared a video with the participants in which they learned about a novel spatial game called the "finding game." In the game, participants were supposed to choose a block out of several options that would match a target block when rotated. Participants were randomly assigned to the structural or control condition (see Table 6). While both conditions described boys as more interested than girls in the finding game, the structural condition highlighted external reasons for this difference. The video said, "Let's talk about why boys are more likely to play the finding game than girls are." In the structural condition, the video emphasized that gender differences were because of how things were set up in the world. Those led to different experiences, which made girls and boys different. Both videos also talked about how when children enter a toy store, the section with boys' toys was more likely to have games like the finding game. In the structural condition, the video highlighted how this structural difference affected children because boys got more chances to play the finding game. However, if the external situation changed, group differences might disappear. Therefore, participants in the structural condition were informed about how structural and environmental factors caused this gender difference.

In contrast, the control condition emphasized internal causes. The video explained that gender differences exist because different kids like different things. The control video also noted that the toy store section with boys' toys was more likely to have games like the finding game. The control video continued to point out internal causes—that boys like to play more with the finding game—and that even if the outside world changed, boys and girls would still be different.

Following the measures, all participants were extensively debriefed, with a particularly detailed debriefing for children in the control condition (Miketta & Friese, 2019). Participants in the control condition learned about the other condition, viewed the structural video, and discussed it with the researcher.

3.3.2 Measures

Attributions about the causes of gender differences were measured with three subscales adapted from Donovan et al. (2019): biological (α = .84), experiential (α = .62, indicating fair reliability), and personal causes (α = .80). Each subscale was measured in three items that related to interests (i.e., *"If girls and boys have different interests, how much is it because of differences inside their bodies, differences in their experiences, or differences in their personal interests?"*), activity choices (*"If girls and boys make different activity choices, how much is it because of differences inside their bodies, differences in their experiences, or differences in their personal activity preferences?"*), and jobs (*"If girls and boys want to do different jobs when they grow up, how much is it because of differences inside their bodies, differences in their experiences, or differences in their personal job preferences?"*). Each of the three causes (i.e., differences inside their bodies, differences in their experiences, or differences in their personal interests) had a sliding bar from 0 to 100, so attributions to each potential cause were measured independently. Each subscale was averaged across the three items (i.e., interest, activity choices, and jobs).

Gender-interest stereotypes about the finding game were measured by the difference score between *"How much would most boys like the finding game if they played it?"* and *"How much would most girls like the finding game if they played it?"* adapted from Master et al. (2021). Both items were measured on a 6-point Likert scale from 1 (*Really not like*) to 6 (*Really like*).

One item served as a manipulation check about the role of experiences by asking, *"Here's a boy who likes the finding game and a girl who does not like the finding game. One kid said that this boy and this girl have different interests because they had different games and experiences growing up. The boy probably got more chances to play the finding game than the girl did. Another kid said that this boy and this girl have different interests because of differences deep inside boys and girls. These differences inside make different kids like different things. Which kid do you think is correct? Are you a little sure, kind of sure, or really sure?*" Participants rated the statement on a 6-point Likert scale from 1 (*First Kid/Really experiences*) to 6 (*Second kid/Really inside*). This item format involving agreement with other children's opinions is often used in developmental psychology studies to ensure children feel that it is socially acceptable to endorse either response (e.g., Heyman & Dweck, 1998).

Interest in the finding game was measured with two items (α = .84, *"How much do you think that you would like to do the finding game?"* and *"How much are you interested in playing the finding game?*), each rated on a 6-point Likert scale from 1 (*Really not interested*) to 6 (*Really interested*).

Future interest development measured the idea of whether participants thought they could develop more of an interest in the finding game in the future. Participants rated their response using two items (α = .61, *"How sure are you that your interest* *in the finding game could change in the future?"* and *"How sure are you that you could develop more of an interest in the finding game?"*), on a 6-point Likert scale from 1 (*Really sure could not*) to 6 (*Really sure could*). Because of low reliability, these two items were analyzed separately when examining condition differences. The inter-item correlation for these two items was r = .44, indicating acceptable internal consistency based on the recommended range of .15-.50 (Clark & Watson, 1995).

The number of activities like the finding game that participants chose to do was included in the study as a measure of behavioral choice. It was scored on a 0 to 3 scale. Participants responded to the question, *"I have some activities like the finding game that you can do at home. I can send them to your family. I can send you zero, one, two, or three of them. How many activities would you like to do?"*

3.3.3 Preregistration

The hypotheses and analyses plan were preregistered on the Open Science Framework at https://osf.io/zwdpyhttps://osf.io/zwdpy/?view_only=420563b571cf4065baabe66ca71fb089. The preregistration specified hypotheses only for interest and choice as the outcome variables, so analyses for all other variables are considered exploratory.

3.4 Results

3.4.1 Preliminary analyses

See Table 7 for correlations between all variables by condition. Current interest was negatively correlated with girls' stereotypes for both the structural explanations condition, $r(107)$ = -.21, p = .028, and control condition, $r(105)$ = -.27, p = .005. Girls who were less likely to endorse stereotypes favoring boys were more interested in the finding game. Girls' beliefs about their future interests (for the structural condition, $r(108)$ = .25, p = .009; for the control condition, $r(108)$ =.21, p = .029) and choice of the number of activities they wanted to do (for structural condition, $r(108)$ = .25, p = .009; for the control condition, $r(108)$ = .40, p < .001) were significantly linked

Table 7 Study 2 Correlations by Condition

Variable	1	2	3	4	5	6	7	8	9
1. Attrib. to biology	–	.04	−.04	.05	.32***	−.03	.06	.04	.03
2. Attrib. to experience	.13	–	.20*	−.03	−.13	.03	−.14	.08	.21*
3. Attrib. to person	.31***	.26**	–	.07	.06	−.06	−.03	.06	.11
4. Gender-interest ster.	.00	−.12	−.06	–	−.05	−.27**	.03	−.20*	−.10
5. Manipulation check	.17	−.11	.10	.01	-	.13	.11	.06	.07
6. Interest	−.08	−.05	−.06	−.21*	−.11	–	.12	.26**	.40***
7. Interest could change	−.12	.07	.13	−.17	−.07	.07	–	.57***	−.14
8. Interest could develop	−.05	.15	.18	−.03	−.22*	.36***	.26**	–	.15
9. Choice	−.07	−.13	-.08	−.01	.11	.25**	−.00	.16	–

Note. The correlations among study variables are presented below the diagonal for the structural condition (N = 110) and above the diagonal for the control condition (N = 110). Attrib. = attribution. Ster. = stereotypes. *p < .05. **p < .01. ***p < .001

to their current interest. Thus, girls who were interested in the finding game also believed that their future interests could develop and were more likely to choose to do the finding game.

3.4.2 Confirmatory analyses from preregistration

See Table 8 for descriptive statistics. For interest, an independent sample t-test was performed to compare interest in the finding game between the structural condition and control condition. Results showed that there was no significant difference between the structural condition, $M = 4.71$, $SD = 1.32$, and the control condition, $M = 4.61$, $SD = 1.27$; $t(218) = 0.59$, $p = .56$, $d = 0.08$.

For choice, according to an independent sample t-test, no significant difference was found in behavioral choice between the two conditions, $t(218) = 0.22$, $p = .83$, $d = 0.03$.

3.4.3 Exploratory analyses (not pre-registered)

Although we did not preregister effects of condition on other measures, we examined them in exploratory analyses. There was no significant condition difference in attributions about gender differences to biological causes, $t(218) = -1.63$, $p = .11$, $d = 0.22$. However, girls in the structural condition were more likely to endorse attributions about gender differences to experiences, $t(218) = 2.48$, $p < .05$, $d = 0.34$, and less likely to endorse attributions of gender differences to personal causes, $t(218) = -2.94$, $p < .01$, $d = 0.40$.

Results showed that there was a significant condition difference in gender-interest stereotypes about the finding game, $t(214) = -4.85$, $p < .001$, $d = 0.66$. Girls in the structural condition were less likely to endorse stereotypes about the finding game that favored boys (see Fig. 2). Although all participants were informed that boys liked the finding game more than girls, gender stereotypes were weakened due to the structural explanation framework, with participants in the control condition perceiving this gender difference as significantly larger compared to participants in the structural condition.

Table 8 Study 2 Descriptive Statistics by Condition

	Structural Explanations		Control Group		Total	
Measures	*M*	*SD*	*M*	*SD*	*M*	*SD*
Attributions to biological causes	26.53	19.72	30.60	17.31	28.56	18.63
Attributions to experiential causes	67.85	14.81	63.08	13.63	65.47	14.40
Attributions to personal causes	69.93	19.38	76.73	14.63	73.33	17.47
Gender-interest stereotypes	.42	1.16	1.29	1.48	.85	1.40
Manipulation check	2.01	1.40	2.14	1.36	2.08	1.38
Interest	4.71	1.32	4.61	1.27	4.66	1.29
Interest could change	5.10	1.20	4.98	1.29	5.04	1.24
Interest could develop	5.23	0.97	4.92	1.26	5.07	1.13
Choice	2.11	.78	2.09	.76	2.10	.76

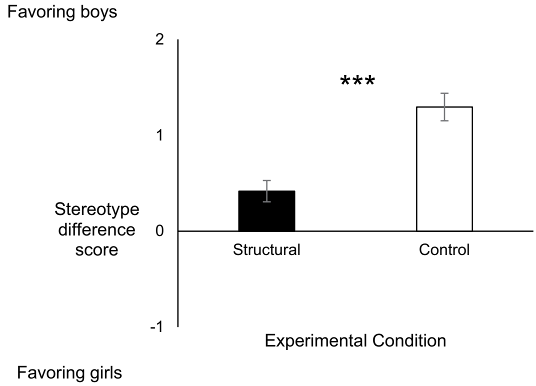

Fig. 2 Structural Explanations Weakened Gender-Interest Stereotypes. The stereotype difference score was calculated by subtracting ratings of how much most girls would like the finding game from how much most boys would like the finding game. Positive values indicate stereotypes favoring boys' interest. ***$p < .001$; $d = 0.66$; error bars are ± s.e

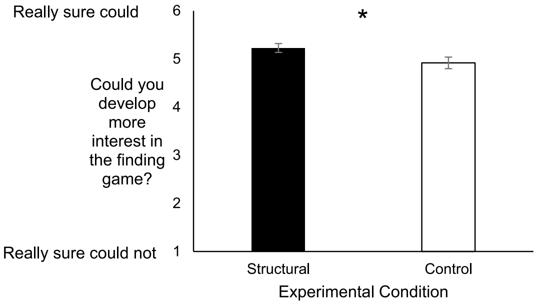

Fig. 3 Structural Explanations Changed Girls' Perceptions of Whether their Interest Could Develop. The y-axis represents how sure girls were that they could develop more of an interest in the finding game, from 1 (really sure could not) to 6 (really sure could). *$p < .05$; $d = 0.28$; error bars are ± s.e

For the manipulation check about the role of experiences, no difference was found between the experimental and control conditions, $t(215) = -.66$, $p = .51$, $d = 0.09$.

The two items measuring future interest development were included separately in analyses. The items measuring whether their interest in the finding game could change showed no significant difference between the structural explanations group and control group, $t(217) = 0.70$, $p = 0.48$, $d = 0.10$. In contrast, participants in the structural condition were more likely to agree that they could develop their interest, $t(218) = 2.04$, $p < .05$, $d = 0.28$, see Fig. 3.

3.5 Discussion

These findings suggest that teaching girls about structural factors is a promising way to change their perceptions of gender differences, although more work is needed to help connect these perceptions to their own STEM motivation. Despite promising findings in a pilot test, we did not find significant effects of structural explanation condition on girls' own interest or behavioral choices to do a novel STEM activity. This suggests that the structural framing may be an important but not sufficient factor to shape girls' own interest. It may also be relevant that girls were not given the opportunity to actually do the activity in this study; they merely learned about it. Future studies should examine whether there may be stronger and longer effects when girls are given actual opportunities to try STEM activities in order to catch and hold their interest (Master et al., 2017). Future studies could also emphasize both within-group comparisons and between-group comparisons by adding descriptions of how girls become more interested in the finding game when they have more opportunities to play it (Amemiya & Bian, 2024). Our manipulation check about the sources of gender differences also suggested that the intervention did not change children's perceptions about gender differences in the finding game, although the means for both conditions showed a belief that gender differences were due to experiences rather than biological differences.

Although the experimental intervention did not change girls' current interest or choices, it did have significant effects on their general attributions about gender differences, stereotypes, and belief that they could develop their interest in the future. We describe three of these effects here. First, while girls in the control condition were more likely to attribute differences to personal interests over experiences, girls in the structural condition made similar attributions to personal interests and experiences. Thus, girls in the structural condition were more likely to recognize the role that experiences play in shaping gender differences. Second, although girls in both conditions were informed that there was a gender difference in interest in the activity, girls in the structural condition perceived that gender difference as significantly smaller than girls in the control condition. This finding aligns with previous research on essentialist perspectives, when differences between categories or groups are perceived as larger than they really are (Master et al., 2012). Finally, there was a condition difference for one of the items about future interest. Girls in the structural condition were significantly more certain that they could develop their own interest in the game in the future. Due to the unreliability of our scale, the small size of this exploratory effect, and a lack of effect on the other item, this finding needs to be replicated. However, this is the kind of belief that might matter over time in efforts to encourage girls to keep an open mind about STEM activities and take advantage of opportunities to develop their interest.

It is also worth noting that there were significant links between attributions, stereotypes, and motivation in both conditions. Perceptions of gender differences and the explanations for those differences have meaningful links to girls' own motivation, although these links may be not causal and/or due to a third variable such as socialized gender roles (Gross et al., 2024).

4 General discussion

Girls are currently underrepresented in many STEM fields, especially computer science and engineering. Gender differences in interest and confidence in these fields emerge early, highlighting the need for interventions to counter negative gender stereotypes and encourage girls' motivation in STEM fields (Master et al., 2025). The two interventions reported here attempted to counter STEM stereotypes and boost girls' STEM motivation by promoting beliefs about the possibility of change. The first intervention (Study 1) examined whether a growth mindset about ability in computer science could promote motivation for computer science for students in Grades 6 and 8. The second intervention (Study 2) was designed to examine whether providing structural explanations for gender differences for a novel activity could promote

11- and 12-year-old girls' motivation for that activity despite the existence of gender stereotypes.

Although neither intervention showed the preregistered predicted effects on motivation, they nevertheless represent promising approaches for future research to pursue. In Study 1, girls in the growth mindset intervention condition reported higher coding interest and identification compared to girls in the control condition. Study 1 also demonstrated that girls' motivation in computer science showed a strong decline between Grades 6 and 8. Future research should examine more fully the causes for this decline in order to prevent this negative trajectory. Correlational findings within Study 1 also showed that mindsets were related to motivation and performance in computer science for students in Grade 8, suggesting that this may be an ideal grade level for future well-powered studies targeting mindsets.

In Study 2, preregistered analyses of the effect of structural explanations on girls' motivation were not significant. However, Study 2 showed (in exploratory analyses) that structural explanations *can* change girls' stereotypes and perceptions of the malleability of interest. These findings make this a promising candidate for future interventions that have the potential to shape the development of girls' interest in real STEM activities.

Although the interventions were not as effective as we had hoped, we believe it is important to publish these findings to avoid "file drawer problems" (in which studies with null results are not published) and so that other researchers may use these openly shared materials to develop more effective future interventions. Null results are challenging to interpret. They may be null due to insufficient contextual understanding by the researchers, materials that are not effective enough, a sample size too small to detect effects, hidden moderators, and/or an incorrect theory of change. Correlational patterns in both studies confirmed the importance of the beliefs we targeted, including mindsets and sources of gender differences. However, it is possible that these beliefs do not have causal effects on motivation. Reverse causality is also possible. Perhaps students who are more motivated see how their effort pays off, which leads to greater growth mindsets. Perhaps girls who feel interested in a STEM activity assume that gender differences must be extrinsic and unstable, because they personally do not feel low interest in STEM. Future studies could aim to improve the intervention materials, including partnering with participating teachers to provide feedback on the materials. Future studies could also attempt to test the reverse causal pattern, or test for hidden moderators of effectiveness.

This pattern of findings also provides some support for our theory of change (Fig. 1). There was more evidence that growth mindsets were beneficial for girls' motivation (see the simple effects of condition for girls on coding interest and identification, although the interaction between condition and gender was not significant for identification) compared to evidence that structural explanations were beneficial. Beliefs that *"I can change"* or *"Others believe I can change"* may be more strongly linked to girls' own motivation than their beliefs that *"my group can change."* Thus, interventions that directly target girls' beliefs about themselves may be more effective than interventions that target more distal beliefs about their group members. More distal interventions (about the group) may need to include personal experiences to make them more impactful for girls' own motivation and choices.

4.1 Limitations and future directions

Although methodological strengths included preregistration and experimental methods, the current studies had several limitations that should be addressed in future research. The growth mindset study (Study 1) showed null results of condition overall, largely nonsignificant interactions between condition and gender (except for interest and belonging), and found that many students failed the attention check. Although our sample size for girls in Grade 8 was large enough to detect an effect on performance even after exclusions, the cell sizes were comparatively small and the failure of the growth mindset manipulation check indicates that these results should not be interpreted as effects of growth mindsets. The particularly small sample size for girls in Grade 8 also undermined our statistical power to test for gender by condition interactions. Future studies should more selectively target this age group to ensure more well-powered intervention studies.

A second limitation is that the dependent measures in Study 1 were not designed to test participants' motivation in the face of struggle or challenge, which is when growth mindsets show the strongest effects according to mindset theory (Burnette et al., 2023). Future studies should carefully design longitudinal intervention studies that can track the development of girls' motivation by providing the intervention near the beginning of the academic year, and then following their motivation as they proceed throughout a computer science course and respond to academic setbacks. Similarly, for the structural intervention, future studies should design longitudinal intervention studies that can track the development of girls' interest in a STEM topic over longer periods of time. Girls in Study 2 did not have the opportunity to engage with the activity before rating their interest. Future studies could also include more authentic and real-world activities, measuring interest after girls have had the chance to actually engage in the activity to examine how these processes might unfold over time. In addition, although both studies were preregistered, many analyses reported here differed from preregistered analyses plans and were accordingly labeled as exploratory. Many of the condition differences from Study 2 were based on these exploratory analyses that should be replicated in future studies.

There were also measurement issues for both studies. Reliability was a problem for measures of mindsets in Study 1 and for the attributions to experiential causes and development of future interest in Study 2. The low internal consistency of these measures indicates potential for a large amount of measurement error and may have reduced effect sizes. Future studies should aim to create and use more reliable scales and measures that can provide a better understanding of how these important beliefs affect girls' motivation in STEM (e.g., Muradoglu et al., 2024). Future studies should also include a pre-test as well as a post-test for these types of interventions to better capture condition effects. This would allow researchers to control for variability in preexisting beliefs about growth mindsets and explanations of gender differences.

We also note that both studies were conducted during the early stages of the COVID pandemic, with Study 2 conducted online using Zoom. This created a technical challenge that could potentially have affected results. However, many developmental research studies show similar effects whether conducted in person or online (Cvencek et al., 2025a, 2025b; Nelson et al., 2021; Zaadnoordijk & Cusack, 2022).

We collected data on how many participants experienced distractions, disruptions, and the Zoom audio/video quality. There was evidence of low or no levels of distractions and disruptions. The audio/video quality was mostly rated by research assistants as "not bad."

4.2 Conclusion

These studies have implications for interventions to help educators counteract negative effects of gender stereotypes and reduce gender gaps in STEM. Many researchers have been working to improve outcomes for girls in STEM (Casad et al., 2018; Perez-Felkner et al., 2025; Prieto-Rodriguez et al., 2020; Rosenzweig & Wigfield, 2016). While some such interventions have been more successful than others, we can learn valuable information even from the interventions that did not generate the expected results. When we remove the barriers that most strongly hold girls back from fulfilling their true potential, we will all benefit. The potential that things can change—that they do not always have to be "this way"—can bring a sense of hope for future possibilities.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11218-026-10294-y.

Acknowledgements We thank participating students and school staff; T. Alexander, J. Amemiya, J. Brown, A. Chavez, A. Cimpian, B. Donovan, O. Duran, C. Dweck, G. Heyman, R. Leshin, C. Lewis, C. Maddox, J. McDonald, F. Oweh, M. Rhodes, B. O. Rivera, M. Rizzo, V. Rodriguez, S. Sriutaisuk, E. Stier, O. Swasey, M. Tennison, N. Vasilyeva; and members of the Stereotypes, Identity, and Belonging and Identity & Academic Motivation Labs.

Author contributions Conceptualization, S.C., A.M., & A.N.M.; Methodology, S.C., A.M., A.N.M. & D.T.; Formal Analysis, K.L., A.M., S.S., & D.T.; Investigation, A.M.; Resources, S.C., A.M., & A.N.M.; Data Curation, A.M.; Writing – Original Draft Preparation, K.L., A.M., S.S., & D.T.; Writing – Review & Editing, S.C., K.L., A.M., A.N.M., S.S., & D.T.; Visualization, K.L., A.M., & S.S.; Supervision, A.M.; Project Administration, A.M.; Funding Acquisition, S.C., A.M., & A.N.M.

Funding National Science Foundation, Grant 1849902 to Dr. Allison Master, Grant 2122488 to Allison Master; Institute of Education Sciences, Grant R305A180167 to Allison Master, Grant R305A200520 to Allison Master

Declarations

Conflict of interest This research was supported by the National Science Foundation (Grants 1849902 and 2122488), and the Institute of Education Sciences, U.S. Department of Education (Grants R305A180167 to the University of Washington and R305A200520 to the University of Houston). The opinions expressed are those of the authors and do not represent views of the Institute of Education Sciences or the U.S. Department of Education. The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results. The authors have no competing interests to declare that are relevant to the content of this article.

References

Amemiya, J., & Bian, L. (2024). Why are there no girls? Increasing children's recognition of structural causes of the gender gap in STEM. *Cognition,245*, Article 105740. https://doi.org/10.1016/j.cognition.2024.105740

Amemiya, J., Mortenson, E., Ahn, S., Walker, C. M., & Heyman, G. D. (2022). Children acknowledge physical constraints less when actors behave stereotypically: Gender stereotypes as a case study. *Child Development,93*(1), 72–83. https://doi.org/10.1111/cdev.13643

Amemiya, J., Mortenson, E., Heyman, G. D., & Walker, C. M. (2023). Thinking structurally: A cognitive framework for understanding how people infer structural causes of inequality. *Perspectives on Psychological Science,18*(2), 259–274. https://doi.org/10.1177/17456916221093593

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2012). "Balancing acts": Elementary school girls' negotiations of femininity, achievement, and science. *Science Education,96*, 967–989. https://doi.org/10.1002/sce.21031

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361–382). The Guilford Press.

Barr, M., Binnie, L., Jacobs, E., Pavlou, K., & West, K. (2024). Understanding the role models that inspire women to study Computing Science. In *Proceedings of the 2024 Conference on United Kingdom & Ireland Computing Education Research*, 1–7. https://doi.org/10.1145/3689535.3689557

Bastian, B., & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology,42*(2), 228–235. https://doi.org/10.1016/j.jesp.2005.03.003

Bigler, R. S. (1999). The use of multicultural curricula and materials to counter racism in children. *Journal of Social Issues,55*(4), 687–705. https://doi.org/10.1111/0022-4537.00142

Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science,16*(3), 162–166. https://doi.org/10.1111/j.1467-8721.2007.00496.x

Bigler, R. S., Spears Brown, C., & Markell, M. (2001). When groups are not created equal: Effects of group status on the formation of intergroup attitudes in children. *Child Development,72*(4), 1151–1162. https://doi.org/10.1111/1467-8624.00339

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development,78*(1), 246–263. https://doi.org/10.1111/j.1467-8624.2007.00995.x

Burnette, J. L., Billingsley, J., Banks, G. C., Knouse, L. E., Hoyt, C. L., Pollack, J. M., & Simon, S. (2023). A systematic review and meta-analysis of growth mindset interventions: For whom, how, and why might such interventions work? *Psychological Bulletin,149*(3–4), 174–205. https://doi.org/10.1037/bul0000368

Burnette, J. L., Hoyt, C. L., Russell, V. M., Lawson, B., Dweck, C. S., & Finkel, E. (2020). A growth mind-set intervention improves interest but not academic performance in the field of computer science. *Social Psychological and Personality Science,11*(1), 107–116. https://doi.org/10.1177/1948550619841631

Carr, P. B., Dweck, C. S., & Pauker, K. (2012). "Prejudiced" behavior without prejudice? Beliefs about the malleability of prejudice affect interracial interactions. *Journal of Personality and Social Psychology,103*(3), 452–471. https://doi.org/10.1037/a0028849

Carroll, J. M., Yeager, D. S., Buontempo, J., Hecht, C., Cimpian, A., Mhatre, P., Muller, C., & Crosnoe, R. (2023). Mindset × context: Schools, classrooms, and the unequal translation of expectations into math achievement. *Monographs of the Society for Research in Child Development,88*(2), 7–109. https://doi.org/10.1111/mono.12471

Casad, B. J., Oyler, D. L., Sullivan, E. T., McClellan, E. M., Tierney, D. N., Anderson, D. A., Greeley, P. A., Fague, M. A., & Flammang, B. J. (2018). Wise psychological interventions to improve gender and racial equality in STEM. *Group Processes & Intergroup Relations,21*(5), 767–787. https://doi.org/10.1177/1368430218767034
Cejka, M. A., & Eagly, A. H. (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality and Social Psychology Bulletin,25*(4), 413–423. https://doi.org/10.1177/0146167299025004002
Cheryan, S., Drury, B. J., & Vichayapai, M. (2013a). Enduring influence of stereotypical coding role models on women's academic aspirations. *Psychology of Women Quarterly,37*(1), 72–79. https://doi.org/10.1177/0361684312459328
Cheryan, S., Lombard, E. J., Hudson, L., Louis, K., Plaut, V. C., & Murphy, M. C. (2020). Double isolation: Identity expression threat predicts greater gender disparities in computer science. *Self and Identity,19*(4), 412–434. https://doi.org/10.1080/15298868.2019.1609576
Cheryan, S., Master, A., & Meltzoff, A. N. (2015). Cultural stereotypes as gatekeepers: Increasing girls' interest in computer science and engineering by diversifying stereotypes. *Frontiers in Psychology,6*, 49. https://doi.org/10.3389/fpsyg.2015.00049
Cheryan, S., Meltzoff, A. N., & Kim, S. (2011). Classrooms matter: The design of virtual classrooms influences gender disparities in computer science classes. *Computers & Education,57*(2), 1825–1835. https://doi.org/10.1016/j.compedu.2011.02.004
Cheryan, S., Plaut, V. C., Handron, C., & Hudson, L. (2013b). The stereotypical computer scientist: Gendered media representations as a barrier to inclusion for women. *Sex Roles: A Journal of Research,69*(1–2), 58–71. https://doi.org/10.1007/s11199-013-0296-x
Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin,143*(1), 1–35. https://doi.org/10.1037/bul0000052
Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment,7*(3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309
Code.org, CSTA, & ECEP Alliance (2022). *2022 state of computer science education: Understanding our national imperative.* Retrieved from https://code.org/assets/advocacy/stateofcs/2022_state_of_cs.pdf
Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (2011). Math–gender stereotypes in elementary school children. *Child Development,82*(3), 766–779. https://doi.org/10.1111/j.1467-8624.2010.01529.x
Cvencek, D., Sanders, E. A., del Río, M. F., Susperreguy, M. I., Strasser, K., Brečić, R., Gaćeša, D., Skala, D., Tomasetto, C., Galdi, S., Cadinu, M., Kapur, M., Passolunghi, M. C., Rueda Ferreira, T. I., Mirisola, A., Mariani, B., & Meltzoff, A. N. (2025a). National disparities favoring males are reflected in girls' implicit associations about gender and academic subjects. *Developmental Psychology,61*(3), 579–593. https://doi.org/10.1037/dev0001797
Cvencek, D., Sanders, E. A., Maddox, C. D., & Meltzoff, A. N. (2025b). New online tools to assess children's implicit social cognition: Replication and generalization of in-person research. *Journal of Educational Psychology,117*(7), 1115–1136. https://doi.org/10.1037/edu0000953
Cyr, E. N., Kroeper, K. M., Bergsieker, H. B., Dennehy, T. C., Logel, C., Steele, J. R., Knasel, R. A., Hartwig, W. T., Shum, P., Reeves, S. L., Dys-Steenbergen, O., Litt, A., Lok, C. B., Ballinger, T., Nam, H., Tse, C., Forest, A. L., Zanna, M., Staub-French, S., … Spencer, S. J. (2024). Girls are good at STEM: Opening minds and providing evidence reduces boys' stereotyping of girls' STEM ability. *Child Development,95*(2), 636–647. https://doi.org/10.1111/cdev.14007
Damore, J. (2017, July). Google's ideological echo chamber. Retrieved from https://assets.documentcloud.org/documents/3914586/Googles-Ideological-Echo-Chamber.pdf
Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to scientific theories affects women's math performance. *Science,314*(5798), 435. https://doi.org/10.1126/science.1131100
Donovan, B. M., Stuhlsatz, M. A., Edelson, D. C., & Buck Bracey, Z. E. (2019). Gendered genetics: How reading about the genetic basis of sex differences in biology textbooks could affect beliefs associated with science gender disparities. *Science Education,103*(4), 719–749. https://doi.org/10.1002/sce.21502
Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.
Dweck, C. S., & Master, A. (2009). Self-theories and motivation: Students' beliefs about intelligence. In K. R. Wenzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 123–140). Routledge/Taylor & Francis Group.
Easterbrook, M. J., & Hadden, I. R. (2021). Tackling educational inequalities with social psychology: Identities, contexts, and interventions. *Social Issues and Policy Review,15*(1), 180–236. https://doi.org/10.1111/sipr.12070

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. https://doi.org/10.1002/(SICI)1098-237X(199901)83

Gál, É., & Kerekes, T. (2025). The relationship between students' and teachers' mindset, self-efficacy and self-esteem: A meta-analysis. *Personality and Individual Differences,246*, Article 113336. https://doi.org/10.1016/j.paid.2025.113336

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195154061.001.0001

Gladstone, J. R., Tallberg, M., Jaxon, J., & Cimpian, A. (2024). What makes a role model motivating for young girls? The effects of the role model's growth versus fixed mindsets about ability and interest. *Journal of Experimental Child Psychology,238*, 1–15. https://doi.org/10.1016/j.jecp.2023.105775

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology,102*(4), 700–717. https://doi.org/10.1037/a0026659

Graham, S., & Taylor, A. Z. (2016). Attribution theory and motivation in school. In K. R. Wentzel & D. B. Miele (Eds.), *Handbook of motivation at school* (pp. 11–33). Routledge.

Gross, E. B., Fine, R. D., Gülgöz, S., Olson, K. R., & Gelman, S. A. (2024). Children's gender essentialism and prejudice: Testing causal links via an experimental manipulation. *Developmental Science,27*(5), Article e13532. https://doi.org/10.1111/desc.13532

Grover, S., Pea, R., & Cooper, S. (2016). Factors influencing computer science learning in middle school. In *Proceedings of the 47th ACM technical symposium on computing science education* (pp. 552–557). New York, NY: ACM. https://doi.org/10.1145/2839509.2844564

Gunderson, E. A., Gripshover, S., Romero, C., Dweck, C. S., Goldin-Meadow, S., & Levine, S. (2013). Parent praise to 1–3 year-olds predicts children's motivational frameworks 5 years later. *Child Development,84*(5), 1526–1541. https://doi.org/10.1111/cdev.12064

Haber, A. S., Kumar, S. C., & Corriveau, K. H. (2022). Boosting children's persistence through scientific storybook reading. *Journal of Cognition and Development,23*(2), 161–172. https://doi.org/10.1080/15248372.2021.1998063

Hecht, C. A., Dweck, C. S., Murphy, M. C., Kroeper, K. M., & Yeager, D. S. (2023). Efficiently exploring the causal role of contextual moderators in behavioral science. *Proceedings of the National Academy of Sciences of the United States of America,120*(1), Article e2216315120. https://doi.org/10.1073/pnas.2216315120

Hecht, C. A., Latham, A. G., Buskirk, R. E., Hansen, D. R., & Yeager, D. S. (2022). Peer-modeled mindsets: An approach to customizing life sciences studying interventions. *CBE—Life Sciences Education,21*(4), Article ar82. https://doi.org/10.1187/cbe.22-07-0143

Heck, R., & Thomas, S. L. (2020). *An introduction to multilevel modeling techniques: MLM and SEM approaches*. Routledge.

Herrmann, S. D., Adelman, R. M., Bodford, J. E., Graudejus, O., Okun, M. A., & Kwan, V. S. (2016). The effects of a female role model on academic performance and persistence of women in STEM courses. *Basic and Applied Social Psychology,38*(5), 258–268. https://doi.org/10.1080/01973533.2016.1209757

Heyman, G. D., & Dweck, C. S. (1998). Children's thinking about traits: Implications for judgments of the self and others. *Child Development,69*(2), 391–403. https://doi.org/10.1111/j.1467-8624.1998.tb06197.x

Heyman, G. D., Dweck, C. S., & Cain, K. M. (1992). Young children's vulnerability to self-blame and helplessness: Relationship to beliefs about goodness. *Child Development,63*(2), 401–415. https://doi.org/10.2307/1131488

Ho, A. K., Roberts, S. O., & Gelman, S. A. (2015). Essentialism and racial bias jointly contribute to the categorization of multiracial individuals. *Psychological Science,26*(10), 1639–1645. https://doi.org/10.1177/0956797615596436

Hong, H.-Y., & Lin-Siegler, X. (2012). How learning about scientists' struggles influences students' interest and learning in physics. *Journal of Educational Psychology,104*(2), 469–484. https://doi.org/10.1037/a0026224

Horwitz, S. R., Shutts, K., & Olson, K. R. (2014). Social class differences produce social group preferences. *Developmental Science,17*(6), 991–1002. https://doi.org/10.1111/desc.12181

Kaya, S., & Karakoc, D. (2022). Math mindsets and academic grit: How are they related to primary math achievement? *European Journal of Science and Mathematics Education,10*(3), 298–309. https://doi.org/10.30935/scimath/11881

Klysing, A. (2020). Exposure to scientific explanations for gender differences influences individuals' personal theories of gender and their evaluations of a discriminatory situation. *Sex Roles: A Journal of Research,82*(5–6), 253–265. https://doi.org/10.1007/s11199-019-01060-w
Lane, K. A., Goh, J. X., & Driver-Linn, E. (2012). Implicit science stereotypes mediate the relationship between gender and academic participation. *Sex Roles,66*, 220–234. https://doi.org/10.1007/s11199-011-0036-z
Law, F., McGuire, L., Winterbottom, M., & Rutland, A. (2021). Children's gender stereotypes in STEM following a one-shot growth mindset intervention in a science museum. *Frontiers in Psychology,12*, Article 641695. https://doi.org/10.3389/fpsyg.2021.641695
Levy, S. R. (1999). Reducing prejudice: Lessons from social-cognitive factors underlying perceiver differences in prejudice. *Journal of Social Issues,55*(4), 745–765. https://doi.org/10.1111/0022-4537.00145
Levy, S. R., & Dweck, C. S. (1999). The impact of children's static versus dynamic conceptions of people on stereotype formation. *Child Development,70*(5), 1163–1180. https://doi.org/10.1111/1467-8624.00085
Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology,74*(6), 1421–1436. https://doi.org/10.1037/0022-3514.74.6.1421
Lin-Siegler, X., Ahn, J. N., Chen, J., Fang, F.-F.A., & Luna-Lucero, M. (2016). Even Einstein struggled: Effects of learning about great scientists' struggles on high school students' motivation to learn science. *Journal of Educational Psychology,108*(3), 314–328. https://doi.org/10.1037/edu0000092
Maltese, A. V., & Tai, R. H. (2010). Eyeballs in the fridge: Sources of early interest in science. *International Journal of Science Education,32*(5), 669–685. https://doi.org/10.1080/09500690902792385
Master, A., Baker, S., Lee, K., & Turcotte, P. (In press). Links between gender stereotypes and gender gaps in STEM motivation in childhood and adolescence. In H. M. G. Watt, E. Mackenzie, B. R. O'Connor, & R. Lazarides (Eds.), *Research Handbook of Gender & STEM*.
Master, A. (2021). Gender stereotypes influence children's STEM motivation. *Child Development Perspectives,15*(3), 203–210. https://doi.org/10.1111/cdep.12424
Master, A., Cheryan, S., & Meltzoff, A. N. (2016). Computing whether she belongs: Stereotypes undermine girls' interest and sense of belonging in computer science. *Journal of Educational Psychology,108*(3), 424–437. https://doi.org/10.1037/edu0000061
Master, A., Cheryan, S., Moscatelli, A., & Meltzoff, A. N. (2017). Programming experience promotes higher STEM motivation among first-grade girls. *Journal of Experimental Child Psychology,160*, 92–106. https://doi.org/10.1016/j.jecp.2017.03.013
Master, A., Markman, E. M., & Dweck, C. S. (2012). Thinking in categories or along a continuum: Consequences for children's social judgments. *Child Development,83*(4), 1145–1163. https://doi.org/10.1111/j.1467-8624.2012.01774.x
Master, A., & Meltzoff, A. N. (2020). Cultural stereotypes and sense of belonging contribute to gender gaps in STEM. *International Journal of Gender, Science, and Technology,12*, 152–198.
Master, A., Meltzoff, A. N., & Cheryan, S. (2021). Gender stereotypes about interests start early and cause gender disparities in computer science and engineering. *Proceedings of the National Academy of Sciences of the United States of America,118*(48), Article e2100030118. https://doi.org/10.1073/pnas.2100030118
Master, A., Meltzoff, A. N., Tang, D., & Cheryan, S. (2025). Divergence in children's gender stereotypes and motivation across STEM fields. *Proceedings of the National Academy of Sciences of the United States of America,122*(18), Article e2408657122. https://doi.org/10.1073/pnas.2408657122
Meyer, M., & Gelman, S. A. (2016). Gender essentialism in children and parents: Implications for the development of gender stereotyping and gender-typed preferences. *Sex Roles,75*(9–10), 409–421. https://doi.org/10.1007/s11199-016-0646-6
Miketta, S., & Friese, M. (2019). Debriefed but still troubled? About the (in)effectiveness of postexperimental debriefings after ego threat. *Journal of Personality and Social Psychology,117*(2), 282–309. https://doi.org/10.1037/pspa0000155
Miller, D. I., Eagly, A. H., & Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology,107*(3), 631–644. https://doi.org/10.1037/edu0000005
Miller, D. I., Lauer, J. E., Tanenbaum, C., & Burr, L. (2024). The development of children's gender stereotypes about STEM and verbal abilities: A preregistered meta-analytic review of 98 studies. *Psychological Bulletin,150*(12), 1363–1396. https://doi.org/10.1037/bul0000456

Mulvey, K. L., Hitti, A., & Killen, M. (2010). The development of stereotyping and exclusion. *Wiley Interdisciplinary Reviews: Cognitive Science,1*(4), 597–606. https://doi.org/10.1002/wcs.66

Muradoglu, M., Porter, T., Trzesniewski, K., & Cimpian, A. (2024). A growth mindset scale for young children (GM-C): Development and validation among children from the United States and South Africa. *PLoS ONE,19*(10), Article e0311205. https://doi.org/10.1371/journal.pone.0311205

Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). Signaling threat: How situational cues affect women in math, science, and engineering settings. *Psychological Science,18*(10), 879–885. https://doi.org/10.1111/j.1467-9280.2007.01995.x

Murphy, M. C., Taylor, V. J., & Steele, C. M. (2024). Stereotype threat: A situated theory of social cognition. In K. Hugenberg, K. Johnson, & D. Carlston (Eds.), *Oxford handbook of social cognition* (2nd ed., pp. 409–441). Oxford University Press.

Nagy, G., Watt, H. M. G., Eccles, J. S., Trautwein, U., Lüdtke, O., & Baumert, J. (2010). The development of students' mathematics self-concept in relation to gender: Different countries, different trajectories? *Journal of Research on Adolescence,20*(2), 482–506. https://doi.org/10.1111/j.1532-7795.2010.00644.x

Nelson, P. M., Scheiber, F., Laughlin, H. M., & Demir-Lira, Ö. E. (2021). Comparing face-to-face and online data collection methods in preterm and full-term children: An exploratory study. *Frontiers in Psychology,12*, Article 733192. https://doi.org/10.3389/fpsyg.2021.733192

O'Keefe, P. A., Horberg, E. J., Dweck, C. S., & Walton, G. M. (2023). A growth-theory-of-interest intervention increases interest in math and science coursework among liberal arts undergraduates. *Journal of Educational Psychology,115*(6), 859–876. https://doi.org/10.1037/edu0000798

Olson, K. R., Dweck, C. S., Spelke, E. S., & Banaji, M. R. (2011). Children's responses to group-based inequalities: Perpetuation and rectification. *Social Cognition,29*(3), 270–287. https://doi.org/10.1521/soco.2011.29.3.270

Ottenbreit-Leftwich, A. T., Kwon, K., Brush, T. A., Karlin, M., Jeon, M., Jantaraweragul, K., Guo, M., Nadir, H., Gok, F., & Bhattacharya, P. (2021). The impact of an issue-centered problem-based learning curriculum on 6th grade girls' understanding of and interest in computer science. *Computers and Education Open,2*, Article 100057. https://doi.org/10.1016/j.caeo.2021.100057

Park, D., Gunderson, E. A., Tsukayama, E., Levine, S. C., & Beilock, S. L. (2016). Young children's motivational frameworks and math achievement: Relation to teacher-reported instructional practices, but not teacher theory of intelligence. *Journal of Educational Psychology,108*(3), 300–313. https://doi.org/10.1037/edu0000064

Passolunghi, M. C., Rueda Ferreira, T. I., & Tomasetto, C. (2014). Math–gender stereotypes and math-related beliefs in childhood and early adolescence. *Learning and Individual Differences,34*, 70–76. https://doi.org/10.1016/j.lindif.2014.05.005

Perez-Felkner, L., Erichsen, K., Li, Y., Chen, J., Hu, S., Ramirez Surmeier, L., & Shore, C. (2025). Computing education interventions to increase gender equity from 2000 to 2020: A systematic literature review. *Review of Educational Research,95*(3), 536–580. https://doi.org/10.3102/00346543241241536

Plaks, J. E., Stroessner, S. J., Dweck, C. S., & Sherman, J. W. (2001). Person theories and attention allocation: Preferences for stereotypic versus counterstereotypic information. *Journal of Personality and Social Psychology,80*(6), 876–893. https://doi.org/10.1037/0022-3514.80.6.876

Plante, I., de la Sablonnière, R., Aronson, J. M., & Théorêt, M. (2013). Gender stereotype endorsement and achievement-related outcomes: The role of competence beliefs and task values. *Contemporary Educational Psychology,38*(3), 225–235. https://doi.org/10.1016/j.cedpsych.2013.03.004

Prieto-Rodriguez, E., Sincock, K., & Blackmore, K. (2020). STEM initiatives matter: Results from a systematic review of secondary school interventions for girls. *International Journal of Science Education,42*(7), 1144–1161. https://doi.org/10.1080/09500693.2020.1749909

Rachmatullah, A., Wiebe, E., Boulden, D., Mott, B., Boyer, K., & Lester, J. (2020). Development and validation of the computer science attitudes scale for middle school students (MG-CS attitudes). *Computers in Human Behavior Reports,2*, Article 100018. https://doi.org/10.1016/j.chbr.2020.100018

Rattan, A., Savani, K., Naidu, N. V. R., & Dweck, C. S. (2012). Can everyone become highly intelligent? Cultural differences in and societal consequences of beliefs about the universal potential for intelligence. *Journal of Personality and Social Psychology,103*(5), 787–803. https://doi.org/10.1037/a0029263

Reeve, J., Basarkod, G., Jang, H. R., Gargurevich, R., Jang, H., & Cheon, S. H. (2025). Specialized purpose of each type of student engagement: A meta-analysis. *Educational Psychology Review,37*(1), 13. https://doi.org/10.1007/s10648-025-09989-z

Reges, S. (2018). Why women don't code. *Quillette.com.* Retrieved from: https://quillette.com/2018/06/19/why-women-dont-code

Régner, I., Steele, J. R., Ambady, N., Thinus-Blanc, C., & Huguet, P. (2014). Our future scientists: A review of stereotype threat in girls from early elementary school to middle school. *Revue Internationale De Psychologie Sociale,27*(3–4), 13–51.

Rodriguez, S. L., & Blaney, J. M. (2021). We're the unicorns in STEM": Understanding how academic and social experiences influence sense of belonging for Latina undergraduate students. *Journal of Diversity in Higher Education,14*(3), 441–455. https://doi.org/10.1037/dhe0000176

Romero, C., Master, A., Paunesku, D., Dweck, C. S., & Gross, J. J. (2014). Academic and emotional functioning in middle school: The role of implicit theories. *Emotion,14*(2), 227–234. https://doi.org/10.1037/a0035490

Rosenzweig, E. Q., & Wigfield, A. (2016). STEM motivation interventions for adolescents: A promising start, but further to go. *Educational Psychologist,51*(2), 146–163. https://doi.org/10.1080/00461520.2016.1154792

Rothbart, M. (2001). Category dynamics and the modification of outgroup stereotypes. In R. Brown & S. L. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup processes* (pp. 45–64). Blackwell Publishing.

Schmidt, F. L. (2011). A theory of sex differences in technical aptitude and some supporting evidence. *Perspectives on Psychological Science,6*(6), 560–573. https://doi.org/10.1177/1745691611419670

Shin, J. E. L., Levy, S. R., & London, B. (2016). Effects of role model exposure on STEM and non-STEM student engagement. *Journal of Applied Social Psychology,46*(7), 410–427. https://doi.org/10.1111/jasp.12371

Smiley, P. A., & Dweck, C. S. (1994). Individual differences in achievement goals among young children. *Child Development,65*(6), 1723–1743. https://doi.org/10.2307/1131290

Snijders, T. A., & Bosker, R. J. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. SAGE.

Spybrook, J., Bloom, H., Congdon, R., Hill, C., Lui, X., Martinez, A., & Raudenbush, S. W. (2005–2013). *Optimal Design Plus Empirical Evidence* (Version 3.01) [Software]. Ann Arbor, MI: University of Michigan, School of Education, Hierarchical Models Project. http://sitemaker.umich.edu/group-based

Steele, J. (2003). Children's gender stereotypes about math: The role of stereotype stratification. *Journal of Applied Social Psychology,33*(12), 2587–2606. https://doi.org/10.1111/j.1559-1816.2003.tb02782.x

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

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education,48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Valian, V. (2014). Interests, gender, and science. *Perspectives on Psychological Science,9*(2), 225–230. https://doi.org/10.1177/1745691613519109

Van Camp, A. R., Gilbert, P. N., & O'Brien, L. T. (2019). Testing the effects of a role model intervention on women's STEM outcomes. *Social Psychology of Education,22*(3), 649–671. https://doi.org/10.1007/s11218-019-09498-2

Vasilyeva, N., Gopnik, A., & Lombrozo, T. (2018). The development of structural thinking about social categories. *Developmental Psychology,54*(9), 1735–1744. https://doi.org/10.1037/dev0000555

Vasilyeva, N., & Lombrozo, T. (2020). Structural thinking about social categories: Evidence from formal explanations, generics, and generalization. *Cognition,204*, Article 104383. https://doi.org/10.1016/j.cognition.2020.104383

Vincent-Ruz, P., & Schunn, C. D. (2018). The nature of science identity and its role as the driver of student choices. *International Journal of STEM Education*. https://doi.org/10.1186/s40594-018-0140-5

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2015). Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering. *Journal of Educational Psychology,107*(2), 468–485. https://doi.org/10.1037/a0037461

Wang, J., Xie, H., & Fisher, J. F. (2011). *Multilevel models: Applications using SAS®*. Walter de Gruyter.

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

Wang, M.-T., Zepeda, C. D., Qin, X., Del Toro, J., & Binning, K. R. (2021). More than growth mindset: Individual and interactive links among socioeconomically disadvantaged adolescents' ability mindsets, metacognitive skills, and math engagement. *Child Development,92*(5), e957–e976. https://doi.org/10.1111/cdev.13560

Wynn, A. T., & Correll, S. J. (2017). Gendered perceptions of cultural and skill alignment in technology companies. *Social Sciences,6*, 1–28. https://doi.org/10.3390/socsci6020045

Xiao, L., Hau, K. T., & Wang, M. D. (2024). Revisiting the usage of alpha in scale evaluation: Effects of scale length and sample size. *Educational Measurement: Issues and Practice,43*(2), 74–81. https://doi.org/10.1111/emip.12604

Yang, X., Naas, R., & Dunham, Y. (2022). Testing the limits of structural thinking about gender. *Developmental Science,25*(2), Article e13169. https://doi.org/10.1111/desc.13169

Yeager, D. S., Carroll, J. M., Buontempo, J., Cimpian, A., Woody, S., Crosnoe, R., Muller, C., Murray, J., Mhatre, P., Kersting, N., Hulleman, C., Kudym, M., Murphy, M., Duckworth, A. L., Walton, G. M., & Dweck, C. S. (2022). Teacher mindsets help explain where a growth-mindset intervention does and doesn't work. *Psychological Science,33*(1), 18–32. https://doi.org/10.1177/09567976211028984

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., Tipton, E., Schneider, B., Hulleman, C. S., Hinojosa, C. P., Paunesku, D., Romero, C., Flint, K., Roberts, A., Trott, J., Iachan, R., Buontempo, J., Yang, S. M., Carvalho, C. M., … Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature,573*(7774), 364–369. https://doi.org/10.1038/s41586-019-1466-y

Zaadnoordijk, L., & Cusack, R. (2022). Online testing in developmental science: A guide to design and implementation. In *Advances in child development and behavior* 62, 93–125. https://doi.org/10.1016/bs.acdb.2022.01.002

Zhang, M. Y., Sullivan, J. N., Markman, E. M., & Roberts, S. O. (2024). Children's structural thinking about social inequities. *Child Development Perspectives,18*(1), 19–25. https://doi.org/10.1111/cdep.12493