

The Effect of Promoting High Expectations on Teachers' Feedback to Students

Yasemein Copur-Gencturk

University of Southern California

Ian Thacker

University of Texas at San Antonio

Joseph R. Cimpian

New York University

Note: This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. The final article will be available, upon publication, via its DOI:

<http://doi.org/10.3102/0013189X261492287>

The Effect of Promoting High Expectations on Teachers' Feedback to Students**Abstract**

Experimental studies show that teachers' feedback expressing high expectations benefits students, but it is unclear whether instructing teachers to emphasize high standards alters their feedback content. We randomly assigned 963 teachers to provide feedback either "as you normally would" or to emphasize "your high standards and that the student has the potential to meet them." Treatment teachers gave more detailed feedback on students' strengths, weaknesses, strategies, effort, and math ability, with less positive ability-related feedback offered to girls and lower effort demanded from White male students. This study is the first to demonstrate that prompting high expectation feedback changes the substance of teacher comments. We discuss implications of promoting high expectations on STEM instruction and promoting equity in STEM.

Introduction

Constructive feedback is a powerful tool for children's social, moral, and intellectual development (Hattie & Timperley, 2007) and steers their learning, motivational, and affective processes (Fong & Schallert, 2023). Feedback helps learners gauge progress, engage in teacher-student dialogue, and view ability as malleable (Fong, & Schallert, 2023; Reeve, 2009; Reeve & Jang, 2006; Senko & Harackiewicz, 2005; Stroet et al., 2013). Students particularly benefit from "wise feedback," feedback that is critical but framed in ways to positively impact how students view themselves (Cohen et al., 1999; Yeager et al., 2014). For example, when teachers frame their feedback in terms of their high expectations and affirm students' potential to meet those expectations, students tend to be more trusting of their teachers' intentions and perform at higher levels in response to critical feedback, with effects being stronger for Black students (Yeager et al., 2014, Experiments 1 and 2). Unsurprisingly, research communicating the importance of tapping into these key psychological processes (Walton & Crum, 2021) is gaining attention, and is widely communicated to practitioners online and in teacher developments (e.g., Greater Good in Education, n.d.; Mindset Works, n.d., University of Connecticut, n.d.).

Although wise feedback can improve student outcomes, little to no research has explored how teachers interpret and apply such feedback strategies in their feedback. Asking teachers to communicate high expectations may increase the likelihood that they do so, but there may be individual differences in how teachers deliver such feedback and how student characteristics influence their feedback. For example, teachers may interpret instructions to use wise feedback in many ways. They may communicate student potential by emphasizing student: strengths, effort, ability, specific mathematical skills, reasoning, or other attributes. Further, student characteristics such as their race and gender, may shape the feedback teachers provide, as

evidence suggests that teachers implicitly judge student ability based on race and gender even when their performance is equivalent (Copur-Gencturk et al., 2020, 2022, 2023) and tend to provide less critical feedback to students of color (Harber et al., 2012, Harber, 2020).

As such, inspired by the research on wise feedback interventions, we wanted to explore whether asking teachers to emphasize their high expectations and students' potential to meet those expectations would alter the content of teacher feedback to hypothetical students, and whether the effects of this intervention differed by the gender and race depicted in images of the students that accompanied their work. In the following section, we describe the elements of teacher feedback examined in this study, informed by prior research and models of feedback, as well as those that might be influenced by prompting teachers to use wise feedback. This is followed by a review of the literature on *changing* teachers' feedback and the role of student race and gender.

Elements of Teacher Feedback

Given that a primary purpose of feedback is to reduce the gap between a learner's current performance on a task and the desired level of performance, the content of teacher feedback plays a critical role in enhancing student learning (e.g., Black & Wiliam, 1998; Hattie & Timperley, 2007; Shute, 2008). In a meta-analysis, Hattie & Timperley (2007) found that, while receiving feedback significantly predicts student performance, its delivery and context are equally important, with largest effects when students received high-information feedback on how to complete tasks in relation to a goal, as compared with more low-information feedback where students received extrinsic rewards, punishments, or praise for task performance. Recent work also found that high-information feedback tends to be effective across a range of student

learning, motivation, and affective outcomes (e.g., Fong & Schallert, 2023; Hattie & Anderman, 2019; Wisniewski et al., 2020).

Our study focuses on six elements of teacher feedback that are identified across multiple models and prior research (including work on wise feedback) as critical components of the feedback content (e.g., Hattie & Timperley, 2007; Panadero & Lipnevich, 2022; Sadler, 2010; Shute, 2008; Yeager et al., 2014): (a) accuracy and specificity in teachers' analysis of student work, (b) attention to the strengths and weaknesses of student performance, (c) specificity of suggestions for improvement, (d) focus level (process, procedures, or student attributes), (e) emphasis on student effort vs ability, and (f) communication of high expectations and potential to meet them (i.e., wise feedback). We discuss the conceptualization and empirical research associated with each element, as well as their potential benefits or detriments to students; and while our study concentrates on teachers and their feedback, we believe it is important to consider how students might respond to these feedback characteristics.

Analysis of Student Work

The accuracy and specificity of teacher feedback analyzing student performance on a particular task are considered key elements of feedback (Hattie & Timperley, 2007; Shute, 2008). When teacher feedback reflects deeper analyses and correct interpretation of student work, students tend to benefit. In an experimental study, Harks et al. (2014) found that more detailed, specific, and accurate “process-oriented” feedback on 9th grade students' math solution strategies (compared to “grade-oriented” feedback) was found to be more effective in promoting their perceptions that the feedback is useful, which then predicted higher achievement and interest in math. A similar quasi-experimental study found that more detailed, specific, and accurate teacher feedback commenting on Singaporean 10th grade students' math misconceptions

helped them strategize a study plan and bolstered their perceptions that the feedback is useful compared to grade-specific feedback (Ng & Kaur, 2014).

Attention to Strengths and Weaknesses of Performance

In addition to providing specific and accurate analyses of student performance, feedback should clearly articulate where the student performed well on a task (strengths of the student work) as well as what still requires improvement (weaknesses; e.g., Shute, 2008). By informing the learner about their strengths and weaknesses, teachers can help students answer three important questions: (a) what are the *goals* students they expected to attain, (b) what *progress* have they made toward achieving these goals, and (c) what *actions* still need to be taken by students to address goals (Hattie & Timperley, 2007; Panadero & Lipnevich, 2022; Rakoczy et al., 2013; Sadler, 2010; Shute, 2008). For this reason, attention to students' strengths and weaknesses represents an important dimension of teacher feedback. At the same time, it is important to distinguish this dimension of feedback from positively or negatively framed feedback. Whereas positive and negative feedback refer to the valence of evaluation, strengths- and weaknesses-focused feedback provides diagnostic information about the quality of specific aspects of student work.

Targeted Suggestions for Improvement

Attending to areas in need of improvement also requires the teacher to offer specific and relevant guidance so that students understand the path to follow to improve their performance (e.g., Black & Wiliam, 1998; Butler & Nisan, 1986; Ramaprasad, 1983). *Suggestions for Improvement* refers to the extent to which teachers' feedback offers targeted suggestions to improve the student's understanding and indicate limitations in student understanding. Research suggests that teacher feedback is more effective when it includes detailed comments about

students' performance and directions for improvement, rather than exclusively providing information about correctness grades (e.g., Ng & Kaur, 2014; Harks et al., 2014; Hattie & Timperley, 2007). In contrast, feedback that is unspecific and offers no suggestions tends to be perceived by students as less useful and constructive (Fong et al., 2021).

Focus Level: Processes, Procedures, or Student Attributes

Hattie and Timperley (2007) highlighted the importance of the focus level of feedback, noting that feedback may be directed at the process level (the strategies students use), the task level (what students did), or the student level (comments about the students themselves). When commenting on student work, mathematics teachers may address how students approached a problem and the soundness of their thinking; what steps and procedures were used; characteristics of the student; or offer generic comments (e.g., “good job”). Process-focused feedback—feedback on procedures, concepts, and effort—benefits students more than person-focused comments on student attributes (Cimpian et al., 2007; Skipper & Douglas, 2012). Overall, feedback on student work, whether positive or negative, fosters adaptive achievement in mathematics (Aumann et al., 2025; Gunderson et al., 2018). Yet classroom feedback varies: some teachers give detailed procedural comments without indicating correctness (Stovner & Klette, 2022), while others note correctness without process feedback (Monteiro et al., 2019).

Ability vs Effort: Comments on Fixed vs Malleable Student Factors

Feedback that comments on individual student characteristics may convey different messages, such as students' fixed capacity to be successful in mathematics (i.e., ability), or malleable processes involved in completing tasks (i.e., effort). Findings generally suggest that students who are praised for their effort (“You must have worked hard on these problems”) tend to show higher task persistence, enjoyment, and performance than individuals who are praised

for their ability (“You must be smart at these problems”; Muller & Dweck, 1998), and feel less helpless in the face of setbacks (Kamins & Dweck, 1999), whether or not the feedback is positive or negative (Aumann et al., 2025; Gunderson et al., 2018).

Communicating High Expectations

The final dimension is framing critical feedback in terms of high expectations. Wise feedback strategies, developed by Cohen, Steele, and Ross (1999), involves setting high standards and assuring students of their potential, signaling respect for their capacity to grow. Yeager et al. (2014) added a brief note to teachers’ essay comments affirming their high expectations, which increased student revision rates and essay quality—especially among Black students who mistrusted school. In follow-up longitudinal work, this intervention also boosted Black and Latino students’ school trust and sense of fairness, reduced disciplinary incidents, and improved four-year college retention (Yeager et al., 2017). Yet most studies used standardized forms of feedback; we know little about the feedback teachers naturally give day-to-day or how interventions might change it.

Research on Changing Teacher Feedback

Most research on changing teacher feedback has taken place in intensive professional development settings and relied on teachers’ self-reports. For example, Fonseca et al. (2015) ran a year-long workshop with 12 secondary teachers in Lisbon, after which participants reported greater confidence and altered feedback practices. Similarly, Gamlem (2015) found that three teachers shifted their beliefs about what constitutes appropriate feedback following a targeted workshop. Single-case research in special education and applied behavior analysis have also created interventions to improve teachers’ use of specific-praise, but results are mixed, effects

are short-term, and only one dimension of feedback was studied (Blood-Pinter et al., 2015; Hawkins & Heflin, 2011; Reinke et al., 2007; Sutherland et al., 2000).

We did find one study that explored how teachers' natural language was modified after professional development. Botello et al. (2025) used an experimental design with five sixth-grade mathematics teachers, finding that those who received professional development on connecting fractions to students' lives were more likely to use motivationally supportive language than a business-as-usual group. We build from this work by scaling up to a larger sample, analyzing multiple dimensions of feedback, and using a controlled setting where all teachers evaluated the same student work.

Teacher Feedback and Student Race and Gender

Some evidence suggests that the feedback teachers provide may be influenced by student characteristics, such as their race and gender. Theory and evidence suggest that some teachers hold biases associating mathematics and science subjects with White/male students, as reflected in their grading, tracking recommendations, implicit associations, and judgments of student potential to succeed in mathematics (see Copur-Gencturk et al, 2020, 2022, 2023, 2024; Denessen et al., 2022; Dee & Gershenson, 2017; Lavy & Sand, 2018; Starck et al., 2020; Thacker et al., 2022). Given consistent findings that teachers' judgments of students are influenced by student factors, we assumed that race and gender bias might also be evident in the written feedback teachers provide to students on their mathematics work.

While we were unable to find research on gender biased feedback, there is some research on teachers' racially biased feedback. Much of the research finds that White teachers provide more *positive*, sympathetic, and lenient feedback on fictitious Black students' substandard essays, though most studies focus on feedback provided on writing assignments (see Harber,

2020 for a review); and the studies that do focus on mathematics-oriented feedback find that teachers provide more *negative* feedback to Black students indicating that the student was incorrect or behaving inappropriately during instruction (Scott et al., 2019). Given this mixed evidence, we aimed to assess whether racial and/or gender biases would be evident in teachers' mathematics feedback.

Summary

Most of the evidence of teacher feedback change relies on self-reports rather than direct observation of teacher feedback or direct observation of teachers' use of specific praise. While there is an existing study that investigates effects on teachers' direct-feedback to students in the classroom resulting from intervention (Botello et al., 2025), that study was smaller in scale, conducted in classrooms where classroom and student characteristics were inconsistent, and did not focus on teachers' written feedback. Further, few studies investigate whether student race and gender influence teachers' mathematics feedback. We aimed to address these gaps by exploring the effects of a simple, "light-touch" intervention prompting teachers to emphasize high expectations, and the role of students' race and gender. To date, no prior experimental research has focused on changing teacher feedback in such a way where student work is held constant across all teachers.

Current Study

No prior experimental research has examined how teachers make sense of and implement high expectations concretely in their own feedback. We addressed this gap by randomly assigning 963 teachers to receive one of two prompts before providing open-ended feedback on student work: the treatment group was asked to "Provide specific feedback to this student, emphasizing your high standards and that the student has the potential to meet those standards,"

while the control group was asked to “Provide specific feedback to this student as you normally would.” We sought to answer the following research questions:

- *How will teachers’ feedback be influenced by prompts to provide feedback that emphasizes the teachers’ high standards and that the student has the potential to meet those standards?*
- *To what extent does the effect of the prompt (i.e., asking teachers to provide feedback emphasizing high expectations and student potential) on teacher feedback differ by student gender and race?*

Methods

Study Context and Sample

This study was part of a larger project designed to investigate the ways gender and racial biases are revealed in teachers’ perceptions and actions, as well as the ways to mitigate such biases. Specifically, the first study examined how teachers’ perceptions of students’ math ability and their recommendations for gifted and special education programs varied by student gender and race (Copur-Gencturk et al., 2022). In this portion of the project, designed as a separate, stand-alone study, we explored how asking teachers to hold high expectations in their feedback altered the content of their feedback and whether the effects of this intervention on teacher feedback differed by student gender and race. The data were collected through an education research company that provides services such as creating a large panel of teachers for a national study and collecting data, as in the case for this study. The teachers were recruited from across the nation and reflected the national profile of the teacher workforce in that the majority of teachers were female (83.6%) and White (68.4%).

Of the 1040 teachers who completed the survey, the analytic sample was restricted to the 963 teachers who provided feedback on all four student solutions. The demographic and school information of the participating teachers as a whole and by the condition to which they were assigned are reported in Table 1. We observed no statistically significant difference between the gender, race, and experience level of the teachers and the group to which they were assigned. Teachers' demographic characteristics were not related to their assigned condition, $\chi^2(2, N=963)=1.33, p=.51$, for gender; $\chi^2(3, N=963)=2.89, p=.41$, for race. Likewise, teachers in the treatment and control groups had similar educational backgrounds. The type of credential teachers held was not related to their condition, $\chi^2(4, N=963)=2.96, p=.56$, or the grade level they were teaching, $\chi^2(2, N=963)=1.40, p=.50$. Teachers in both conditions had similar levels of mathematics teaching experience, for a difference of 0.38 years, $t(961)=0.67, p=.51$.

[Table 1 Here]

Experimental Design

The goal of this study was to explore the impact of a treatment that asked teachers to hold high expectations in their feedback to students who had not solved the mathematics problems fully correctly. Each teacher was asked to provide open-ended feedback on two incorrect and two partially correct student solutions. For all four solutions, there was room for teachers to express high expectations of students who were not excelling and to provide guidance to improve their performance. The student solutions were obtained from middle school students who answered mathematics problems used on the National Assessment of Educational Progress, a national assessment.

Student work was presented to teachers in screenshots of online learning classrooms, where individual students appeared to share their solutions (see Figure 1). The images of these

students were created by holding individual meetings with volunteer middle-grade students of different races and genders. We randomly assigned these images to each of the mathematics solutions used in the experiment so that we could control the potential variation in teachers' feedback, depending on students' race and gender.

[Figure 1 Here]

Teachers were randomly assigned to one of two conditions: (1) a treatment, in which the teachers were asked to provide feedback emphasizing their high standards and students' potential to meet those standards, or (2) the control condition, in which teachers were asked to provide feedback to students in the way they normally would. Teachers in both conditions were asked to evaluate the same student solutions, with the only difference being the race- and gender-specific images that were randomly assigned to these solutions. After they had completed providing feedback, the teachers were asked to provide information regarding their demographic and educational background.

Coding Teachers' Responses

We coded teachers' feedback prior to conducting statistical analyses, blind to student gender, race, and condition. Informed by prior research and multiple conceptualizations of feedback, we first identified the feedback elements (i.e., coding areas) used to code the data (Busch et al., 2015; Ellegaard et al., 2018; Evans, 2013; Guo & Wei, 2019; Hattie & Timperley, 2007; Kim, 2016; Kluger & DeNisi, 1996; Knight et al., 2021; Rakoczy et al., 2013; Schleppenbach et al., 2007; Shute, 2008). This process guided the development of the initial coding scheme, including definitions of categories to capture distinct aspects of each feedback element (Panadero & Lipnevich, 2022). Following an iterative, data-informed approach (Saldaña, 2015), the first author and two coders met to apply the initial codes to a randomly

selected subset of responses, to refine and clarify the coding categories. After establishing initial agreement and finalizing the coding categories, two raters independently coded each response, assigning a code for each coding area (see Table 2 for brief descriptions and sample responses for each category across coding areas). The level of agreement, on average (i.e., kappa statistic of .65), ranged from .55 to .79 (Landis & Koch, 1977). The exact percent agreement between the raters was on average 84%, ranging from 80% to 92% agreement for each coding category. Any disagreements were resolved by discussion (Harry et al., 2005), and a final code was assigned.

[Table 2 Here]

Condition. To capture the impact of the treatment, a binary variable was created to indicate the condition to which teachers were assigned, with the control condition being the reference category.

Student Gender and Race. We created binary-coded variables to capture the gender and race of the students in the images, with White and male as the referent categories.

Student Solutions. A set of variables was created to indicate which student solution teachers provided feedback on, so the role of specific student solutions in observed outcomes could be considered in the reported analyses.

Analytic Approach

For both research questions, we conducted a multinomial logistic regression analysis for each of the coding areas as the categories within each coding area were nominal. Each analysis included the following variables: the gender and race of the student images, the student solutions, and the condition to which teachers were assigned, and the interaction terms among gender, race, and condition. To address the first research question, we reported the overall main effects of the treatment. To address the second research question (i.e., whether the impact of the intervention

on teacher feedback differed by student gender and race), we examined treatment effects for specific gender and race combinations (i.e., White girls, Black girls, White boys, and Black boys). For all analyses, we clustered the error within teachers because each teacher provided feedback to four student solutions. We used robust standard error estimations in the regression analyses, which do not rely on assumptions of the homoskedasticity of error. Because we conducted a total of six regression analyses for each area of teacher feedback per research question, we adjusted the p-value to reduce the artificially increased discovery rate through multiple analyses by applying the Bonferroni correction.

Preliminary Analysis

As a preliminary check, we examined whether teachers in the treatment group literally wrote about high standards and student potential. Only 2.4% of treatment responses explicitly mentioned these ideas, consistent with a well-documented pattern in teacher professional development research: teachers' uptake of interventions often varies (e.g., Fink & Stoll, 2005; Gregoire, 2003; März & Kelchtermans, 2013), as they process and apply information in their own ways. Treatment teachers provided lengthier feedback than control teachers (words: $M=35.90$, $SD=25.10$ vs. $M=28.47$, $SD=21.47$; characters: $M=187.57$, $SD=127.75$ vs. $M=145.33$, $SD=107.88$), with both differences statistically significant (word count: $t(3850)=9.89$; character count: $t(3850)=11.12$; both $p<.0001$). Our subsequent analyses therefore focus on how the prompt shifted the content of teachers' feedback more broadly.

Results

Promoting High Expectations Increased Teachers' Attention to Both Strengths and

Weaknesses of Student Solutions

Across all student solutions, teachers in the treatment condition did not comment on either the strengths or weakness of the student solutions 8.6% of the time, while this ratio was 16.0% for teachers in the control group (Table 3). This difference of 7.4 percentage points corresponds to a relative ratio (RR) of 0.54, meaning that teachers in the treatment condition were about half as likely to not comment on either strengths or weakness of the solution ($p < .0001$). Teachers in the treatment condition commented on what worked and what did not in student approach (i.e., both the strengths and weaknesses of students' approach) 38.6% of the time, whereas teachers in the control group did so 26.4% of the time ($RR = 1.46$, $p < .001$). While teachers in the treatment group did not focus on only the weaknesses of student solutions compared with teachers in the control group (35.8% versus 40.0%), this result was not robust to the Bonferroni correction ($p = .01$). We found no difference in how often teachers in the treatment and control groups commented on only the strengths of students' solutions.

[Table 3 Here]

Promoting High Expectations Increased Teachers' Focus on Student Attributes

Teachers in the treatment group shifted their attention to students rather than their work 4.2% of the time, while this percentage was 0.6% for control group teachers ($RR = 7.00$, $p < .0001$). We found no difference between these two groups of teachers in the extent to which they focused on the steps the students took to answer the problem (i.e., what the students did) or the process involved in the students' strategies (i.e., how the students solved the problem).

Promoting High Expectations Increased Teachers' Focus on Students' Math Ability and Effort

Teachers in the treatment group commented on student math ability negatively 2.7% of

the time whereas teachers in the control group did so 0.5% of the time ($RR=5.40, p=.001$). The treatment also increased teachers' praise of student math ability. Teachers in the treatment group commented positively on students' math ability 8.1% of the time, while only 2.1% of teachers in the control group did so ($RR=3.86, p<.0001$).

Despite the fact that the treatment had no impact on teachers' positive praise of students' effort, teachers in the treatment condition commented that students needed to put more effort into their work 2.8% of the time while teachers in the control group did so only 0.5% ($RR=5.60, p<.001$).

Promoting High Expectations Had No Effect on Teachers' Analysis of Student Solutions

Asking teachers to emphasize their high expectations for students as well as the students' potential to meet those expectations had no detectable effect on teachers' evaluations of the correctness of students' solutions. There were no statistically significant treatment effects for any of the student gender-race groups (see Table S4 in the online Supplementary Materials) after the Bonferroni correction that adjusts for multiple comparisons (i.e., $\alpha=.05/6=.008$).

Promoting High Expectations Had No Impact on Teachers' Targeted Suggestions to Improve Student Understanding

Finally, we explored the ways in which teachers provided guidance to improve students' learning in the areas where students needed guidance. We grouped teachers' responses into five categories depending on the specificity of their suggestions and the extent to which they clarified in their feedback how these suggestions might overcome the struggles students demonstrated in their solutions (see Table S1 in the online Supplementary Materials for further information on these categories). Our results suggested that asking teachers to emphasize their high expectations

did not affect the frequency with which teachers offered a strategy that could help students overcome the struggle.

Inducing High Expectations Increased Negative Ability Feedback for White and Black Girls, Reduced Neutral Effort Feedback for Black Students, and Increased Negative Effort Feedback for All Other than White Males

The overall impact on teacher feedback of promoting high expectations did not differ by student gender or race (see the online Supplementary Materials for detailed results for each coding area). One exception was observed in teachers' attention to the strengths and weaknesses of student solutions. As noted previously, although the intervention reduced the proportion of feedback that attended to neither strengths nor weaknesses overall, this pattern did not hold for White boys after the Bonferroni correction ($RR=0.69, p=.02$). Specifically, teachers in both conditions were similarly unlikely to attend to either strengths or weaknesses when responding to solutions accompanied by images of White boys (Figure 2).

Although teacher feedback rarely included negative comments about ability or effort, differential treatment effects by student gender and race emerged in these categories. Teachers in the treatment group provided more negative feedback to girls: both White girls and Black girls received more negative comments about their math ability from teachers in the treatment condition ($RR=12.3, p=.003$, and $RR=5.7, p=.002$, respectively; Figure 2). Another differential effect of the treatment was observed in the frequency of "no or mixed" comments on student effort, which appeared less often for Black boys and girls in the treatment group ($RR=0.95, p<.001$, and $RR=0.96, p=.002$, respectively; see Figure 2). Finally, White boys were the only group that did not receive significantly increased negative feedback about effort from teachers in the treatment condition ($RR=2.44, p=.08$; Figure 2).

[Figure 2 Here]

Robustness Check

We conducted two additional sets of analyses to check the robustness of our findings regarding the directness of teacher feedback. Although third-person feedback is recommended when providing negative feedback (e.g., Brinko, 1993), and is frequently used in teacher education contexts, it may differ from direct, student-addressed feedback in ways that raise validity concerns. To address this issue, a coder who is not part of the authorship team was asked to first identify feedback written entirely in the third person versus feedback addressed directly to the student, and second, to expand the search to identify feedback containing a mix of third-person and student-focused language (see online Supplementary Text 1 for examples). The first coding approach identified 381 third-person feedback responses from 126 teachers (see the results for after removing these cases in the “third-person feedback removed” columns in online Tables S1–S5 in the Supplementary Materials; analytic sample $N=894$ teachers). The second approach identified 426 feedback responses from 151 teachers (see the results for after removing these cases in the “direct feedback” columns in Tables S1–S5; $N=886$ teachers). Overall, the significance of the treatment effects across the feedback dimensions remained similar, with the exception of the student-focused feedback, which became not significant, and some differential treatment effects for gender and race interactions in the “negative ability” and “process” categories (see the online supplementary findings for details). Given the minimal differences and the common use of third-person feedback in teacher education contexts, we focus our discussion on the unadjusted results.

Discussion

We investigated whether prompting teachers to emphasize high expectations and student potential would alter their feedback, and if feedback was different depending on students' race or gender. Compared to a control group, teachers prompted to emphasize high expectations were less likely to provide vague feedback and more likely to provide detailed information about students' strengths, weaknesses, and provide feedback about students and their effort and ability. Regarding the latter, we also found that teachers offered less positive ability feedback to girls and demanded more effort from all students other than White boys. Teacher feedback did not significantly differ in the frequency that they emphasized mathematical procedures or strategies. As such, our findings suggest that a simple treatment reminding teachers to emphasize their high expectations and students' potential to meet those expectations can bolster the amount of targeted information teachers provide, mostly in ways that benefit students, but not always. In what follows, we discuss in more detail the dimensions of feedback that were shifted by wise feedback prompts, the potential impacts these may have on students, and feedback differences based on student race and gender.

Dimensions of Feedback Change and Potential Implications for Students

Interestingly, encouraging teachers to communicate their high expectations shaped how they provided feedback, but in some unexpected ways. The prompt included no explicit information about several of the dimensions of teacher feedback that were observed to be influenced by the treatment. Teachers were not told explicitly to alter aspects of the feedback related to specificity, strengths/weaknesses, or effort/ability; nevertheless, being prompted to emphasize their high expectations had such effects, for better or for worse. On the one hand, fewer generic suggestions and more targeted feedback that helps students understand expectations and goals, progress towards goals, and actions needed to meet goals (i.e., their

strengths and weaknesses) can be beneficial for student learning and can provide information about the students' level of competence (Hattie & Timperley, 2007; Shute, 2008; Wisniewski et al., 2020). Indeed, specific analysis of student work and their solution processes (Harks et al., 2014; Ng & Kaur, 2014), attention to student strengths and weaknesses (Hattie & Timperley, 2007), and effort emphases (Kamins & Dweck, 1999; Lam et al. 2008; Muller & Dweck, 1998) tends to bolster student motivation and/or performance, depending on students' ability beliefs and goal orientations.

On the other hand, emphasizing students' ability in feedback is often considered counter-productive and can undermine students' motivation when they encounter setbacks (Dweck & Leggett, 1988; Fong et al., 2021). Indeed, feedback directed at student effort, positive or negative, generally leads to more adaptive outcomes (Yeager & Dweck, 2012). And, when inconsistent praise is given (e.g., sometimes process-focused, sometimes person-focused), any amount of person-focused feedback has negative effects on students' motivations and behavior, through implying that an underlying trait affects outcomes (Zentall & Morris, 2010). However, it is possible that asking teachers to emphasize high expectations inadvertently primed teachers to think more about student characteristics, aligning with Brophy's (1981) observation that teachers' beliefs about student ability influence their feedback practices. This discrepancy between intended and actual outcomes highlights the complexity of implementing such feedback strategies. While the goal is to communicate high standards and belief in students' potential to meet them, teachers may need more specific guidance on how to do so without resorting to ability-focused language. We encourage researchers to replicate our study to investigate whether these findings might hold more generally, investigate strategies for diminishing teachers' emphasis on ability and enhancing the emphasis on effort and targeted corrective feedback. Our

study made use of a “light-touch” intervention to serve as an initial investigation of the kinds of feedback teachers provide, and whether such feedback might be shifted with prompts. However, future work might engage teachers in more involved learning experiences that prepare them with the theory and reasoning behind these feedback practices, supporting evidence of its effectiveness, examples, and activities to model motivationally supportive feedback in the classroom (see, e.g., Botello et al., 2025).

Despite our findings that teachers in the treatment group provided feedback that may have mixed impacts on students, our findings contribute to the research on teacher education in several ways. First, our findings demonstrate that specific properties of teacher feedback can be changed with intervention. While some prior research has investigated changes to teachers’ feedback, specifically by looking at changes in their use of specific praise (e.g., Blood-Pinter et al., 2015; Reinke et al., 2007; Sutherland et al., 2000) or self-reported learning outcomes (e.g., Fonseca et al., 2015; Gamlem, 2015), our study is the first larger-scale study to openly explore multiple dimensions of teachers’ natural feedback, and the influence that a prompt to emphasize high expectations has on those dimensions. Future research might implement our feedback rubrics in other research contexts and/or train large-language models for evaluating teacher feedback based on our open-access data (see online Supplemental Materials). Recent efforts to build large-language models trained on one-on-one tutoring sessions and classroom interactions have shown to have benefits for teachers and student learning and persistence (Demszky et al., 2024; Malamut et al., 2025), yet to our knowledge, no tools have been developed using teachers’ written feedback on student work. Moreover, future research might explore additional dimensions of teacher feedback as they pertain to a wider variety of topics in STEM or other

domains and identify interventions that encourage teachers' use of supportive and effective feedback.

Second, this study contributes to research on interventions that emphasize high expectations. Our findings suggest that stressing high expectations, which tends to be effective in controlled experiments with students (e.g., Yeager et al., 2014), may have unintended consequences when interpreted and implemented by practitioners. Research on similar psychosocial interventions is widely communicated to practitioners (e.g., Greater Good in Education, n.d.; Mindset Works, n.d.) with little investigation of how such messages shape practice. Future research might investigate how to modify such approaches to curtail teachers' emphasis on ability while enhancing the more beneficial feedback shifts and examine downstream effects of these feedback dimensions on student outcomes.

Student Race and Gender and Teacher Feedback

In addition to finding overall effects of the intervention on teacher feedback, we found small but significant differences in teachers' responses based on student race and gender. First, we found that teachers in the treatment condition more frequently provided negative ability feedback for girls. This finding is consistent with a pattern found in the literature: teachers tend to attribute the low performance of boys to lack of effort, while for girls, a lack of mathematical ability (e.g., Copur-Gencturk et al., 2020; 2022; 2023; Fennema et al., 1998; Tiedemann, 2000; Thacker et al., 2022). While this prior research has found that teachers make gender-based tracking decisions and judgments of student potential in mathematics, our findings suggest that such judgments may also manifest in teachers' feedback—they provided more explicit negative feedback to girls about their ability for incorrect work. Why these biases surfaced more often in the intervention condition deserves further investigation, but one explanation may be that the

prompts asking teachers to reflect on student potential may have activated biases among those who might resist the notion that girls have such potential (see, e.g., Dersch et al., 2026).

Second, regarding effort-based feedback and attention to strengths and weaknesses, we found that teachers in the treatment condition provided less neutral or no effort feedback for Black students, and more negative effort feedback and attention to strengths and weaknesses to all students other than White male students. These findings suggest that, unlike the research focused on teachers' critiques of student writing assignments (Harber, 2020), teachers in our study did not provide more lenient feedback to Black students compared to White students. Rather, teachers' effort critiques and attention to strengths/weaknesses were slightly less common when teachers were commenting on White male students work when asked to emphasize high expectations. One explanation is that prompting teachers to use wise feedback strategies may trigger them to attend to underrepresented students in STEM, who tend to benefit most from wise feedback (Yeager et al., 2014), and direct this attention toward these students' effort, strengths, and weaknesses.

Limitations

We would like to acknowledge some of this study's limitations. First, this design required teachers to provide feedback to students they did not know and to respond to tasks they might not have selected themselves. To mitigate these concerns, we asked teachers to provide feedback as they naturally would, selected problems from a national assessment, and provided authentic student solutions. Notably, the feedback patterns observed here are consistent with classroom observation research on teacher feedback focus levels (Monteiro et al., 2019; Smit et al., 2024; Stovner & Klette, 2022), providing some convergent validity for our findings. Second, the treatment studied was a very "light-touch" intervention, and future research should explore

whether the effects of being asked to emphasize high expectations could be guided through a more sustained and involved intervention, such as a professional development session or series on how to provide effective feedback.

Conclusion

Communicating high expectations is a promising strategy for improving the quality and receptiveness of teacher feedback. However, teacher training programs that promote high expectations should also help teachers understand how to emphasize student potential without defaulting to ability-focused language, and how to deliver equitable feedback across student gender and race groups.

References

- Aumann, L., Puca, R. M., & Gasteiger, H. (2025). Effects of early childhood teacher knowledge and feedback on children's early mathematical development in German kindergartens. *Early Childhood Education Journal*, 1-11.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), 7-74.
- Blood-Pinter, E. B., East, A. & Thrush, N. (2015). Effects of a video-feedback intervention on teachers' use of praise. *Education and Treatment of Children*, 38, 451–472
- Botello, M., Dyson, N., Rutherford, T., & Jordan, N. C. (2025). Teachers' Use of Motivational Messages While Teaching Fractions. *The Elementary School Journal*, 125(4), 1–24.
- Brinko, K.T. (1993). The practice of giving feedback to improve teaching: What is effective? *The Journal of Higher Education*, 64(5), 574-593.
- Brophy, J. (1981). Teacher praise: A functional analysis. *Review of Educational Research*, 51(1), 5-32.
- Busch, J., Barzel, B., & Leuders, T. (2015). Promoting secondary teachers' diagnostic competence with respect to functions: development of a scalable unit in Continuous Professional Development. *ZDM*. 47, 53-64.
- Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology*, 78(3), 210.
- Cimpian, A., Arce, H. M. C., Markman, E. M., & Dweck, C. S. (2007). Subtle linguistic cues affect children's motivation. *Psychological Science*, 18(4), 314-316.

- Cohen, G. L., Steele, C. M. & Ross, L. D. (1999). The mentor's dilemma: Providing critical feedback across the racial divide. *Personality and Social Psychology Bulletin* 25(10), 1302–1318.
- Copur-Gencturk, Y., Cimpian, J. R., Lubienski, S. T., & Thacker, I. (2020). Teachers' bias against the mathematical ability of female, Black, and Hispanic students. *Educational Researcher*, 49(1), 30–43.
- Copur-Gencturk, Y., Thacker, I., & Cimpian, J. R. (2022). Teacher bias in the virtual classroom. *Computers & Education*, 191, 104627. <https://doi.org/10.1016/j.compedu.2022.104627>
- Copur-Gencturk, Y., Thacker, I., & Cimpian, J. R. (2023). Teachers' race and gender biases and the moderating effects of their beliefs and dispositions. *International Journal of STEM Education*, 10(1), 31. <https://doi.org/10.1186/s40594-023-00420-z>
- Copur-Gencturk, Y., Thacker, I., & Cimpian, J. R. (2024). An exploratory experiment investigating teachers' attributional race and gender bias and the moderating effects of personal experience of racial discrimination. *Contemporary educational psychology*, 79, 102317. <https://doi.org/https://doi.org/10.1016/j.cedpsych.2024.102317>
- Dee, T., & Gershenson, S. (2017). *Unconscious Bias in the Classroom: Evidence and Opportunities*. Mountain View, CA: Google Inc. Retrieved from <https://goo.gl/O6Btqi>.
- Demszky, D., Liu, J., Hill, H. C., Jurafsky, D., & Piech, C. (2024). Can automated feedback improve teachers' uptake of student ideas? Evidence from a randomized controlled trial

- in a large-scale online course. *Educational Evaluation and Policy Analysis*, 46(3), 483-505.
- Denessen, E., Hornstra, L., Van Den Bergh, L., & Bijlstra, G. (2022). Implicit measures of teachers' attitudes and stereotypes, and their effects on teacher practice and student outcomes: A review. *Learning and Instruction*, 78, 101437.
- Dersch, A. S., Thacker, I., & Eitel, A. (2026). Should we Tell Teachers That They Are (In) Correct? How Personalized Refutation Text Promotes or Hampers Conceptual Change. *Educational Psychology Review*, 38(1), 17.
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95(2), 256.
- Ellegaard, M., Damsgaard, L., Bruun, J., & Johannsen, B. F. (2018). Patterns in the form of formative feedback and student response. *Assessment & Evaluation in Higher Education*, 43(5), 727-744.
- Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research*, 83(1), 70-120.
- Fennema, E., Carpenter, T. P., Jacobs, V. R., Franke, M. L., & Levi, L. W. (1998). A longitudinal study of gender differences in young children's mathematical thinking. *Educational Researcher*, 27(5), 6-11.
- Fink, D., & Stoll, L. (2005). Educational Change: Easier Said than Done. In: Hargreaves, A. (Eds) Extending Educational Change. Springer, Dordrecht. https://doi.org/10.1007/1-4020-4453-4_2.

- Fong, C. J. & Schallert, D. L. (2023). “Feedback to the future”: Advancing motivational and emotional perspectives in feedback research. *Educational Psychologist*, 58(3), 146–161.
- Fong, C. J., Schallert, D. L., Williams, K. M., Williamson, Z. H., Lin, S., Kim, Y. W., & Chen, L. H. (2021). Making feedback constructive: the interplay of undergraduates' motivation with perceptions of feedback specificity and friendliness. *Educational Psychology*, 41(10), 1241-1259.
- Fonseca, J., Carvalho, C., Conboy, J., Valente, M. O., Gama, A. P., Salema, M. H., & Fiuza, E. (2015). Changing teachers' feedback practices: A workshop challenge. *Australian Journal of Teacher Education*, 40(8), 59-82.
- Gamlem, S. M. (2015). Feedback to support learning: changes in teachers’ practice and beliefs. *Teacher Development* 19, 461–482.
- Greater Good in Education. (n.d.) *Giving wise feedback*. Retrieved, September, 10, 2026.
<https://ggie.berkeley.edu/practice/giving-wise-feedback/>
- Gregoire, M. (2003). Is it a challenge or a threat? A dual-process model of teachers' cognition and appraisal processes during conceptual change. *Educational Psychology Review*, 15(2), 147-179.
- Gunderson, E. A., Sorhagen, N. S., Gripshover, S. J., Dweck, C. S., Goldin-Meadow, S., & Levine, S. C. (2018). Parent praise to toddlers predicts fourth grade academic achievement via children’s incremental mindsets. *Developmental Psychology*, 54(3), 397.
- Guo, W., & Wei, J. (2019). Teacher Feedback and Students’ Self-regulated Learning in Mathematics: A Study of Chinese Secondary Students: G. Wenjuan, W. Jun. *The Asia-Pacific Education Researcher*, 28(3), 265-275.

- Harber, K. D. (2020). The positive feedback bias. In J. Aronson & E. Aronson (Eds.), *The social animal* (11th ed., pp. 475–494). Worth Publishers.
- Harber, K. D., Gorman, J. L., Gengaro, F. P., Butisingh, S., Tsang, W., & Ouellette, R. (2012). Students' race and teachers' social support affect the positive feedback bias in public schools. *Journal of Educational Psychology*, 104(4), 1149.
- Harks, B., Rakoczy, K., Hattie, J., Besser, M., & Klieme, E. (2014). The effects of feedback on achievement, interest and self-evaluation: the role of feedback's perceived usefulness. *Educational Psychology*, 34(3), 269-290.
- Harry, B., Sturges, K. M. & Klingner, J. K. (2005). Mapping the process: An exemplar of process and challenge in grounded theory analysis. *Educational Researcher* 34, 3–13.
- Hattie, J. & Anderman, E. M. (2019). *Visible learning guide to student achievement: Schools edition*. Routledge.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Hawkins, S. M., & Heflin, L. J. (2011). Increasing secondary teachers' behavior-specific praise using a video self-modeling and visual performance feedback intervention. *Journal of Positive Behavior Interventions*, 13(2), 97-108.
- Kamins, M. L., & Dweck, C. S. (1999). Person versus process praise and criticism: implications for contingent self-worth and coping. *Developmental Psychology*, 35(3), 835.
- Kim, Y. (2016). Interview prompts to uncover mathematical knowledge for teaching: Focus on providing written feedback. *The Mathematics Enthusiast*, 13(1), 71-92.

- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254.
- Knight, S. K., Greenberger, S. W., & McNaughton, M. E. (2021). An interdisciplinary perspective: The value that instructors place on giving written feedback. *Active Learning in Higher Education*, 22(2), 115-128.
- Lam, S. F., Yim, P. S., & Ng, Y. L. (2008). Is effort praise motivational? The role of beliefs in the effort–ability relationship. *Contemporary Educational Psychology*, 33(4), 694-710.
- Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159–174.
- Lavy, V., & Sand, E. (2018). On the origins of gender gaps in human capital: Short-and long-term consequences of teachers' biases. *Journal of Public Economics*, 167, 263-279.
- März, V., & Kelchtermans, G. (2013). Sense-making and structure in teachers' reception of educational reform. A case study on statistics in the mathematics curriculum. *Teaching and Teacher Education*, 29, 13-24.
- Malamut, J., Demszky, D., Bywater, C., Reinhart M., & Hill, H. C. (2025). *Facilitating Evidence-Based Instructional Coaching With Automated Feedback on Teacher Discourse*. (EdWorkingPaper: 25-1298). Annenberg Institute at Brown University: <https://doi.org/10.26300/xx9z-8f27>
- MindsetWorks. (n.d). *Programs that motivate students and teachers*. Retrieved, September, 10, 2026. <https://www.mindsetworks.com/>

- Monteiro, V., Mata, L., Santos, N., Sanches, C., & Gomes, M. (2019). Classroom talk: The ubiquity of feedback. *Frontiers in Education* 4(140), 1–18.
- Muller, C. & Dweck, C. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, 75, 33–52.
- Ng, S. K. B., & Kaur, B. (2014). Using enhanced feedback to improve the learning of mathematics. *The Mathematics Educator*, 15(2), 101-119.
- Panadero, E., & Lipnevich, A. A. (2022). A review of feedback models and typologies: Towards an integrative model of feedback elements. *Educational Research Review*, 35, 100416.
- Rakoczy, K., Harks, B., Klieme, E., Blum, W., & Hochweber, J. (2013). Written feedback in mathematics: Mediated by students' perception, moderated by goal orientation. *Learning and Instruction*, 27, 63-73.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral science*, 28(1), 4-13.
- Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist*, 44(3), 159–175.
- Reeve, J. & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, 98(1), 209.
- Reinke, W. M., Lewis-Palmer, T. & Martin, E. (2007). The effect of visual performance feedback on teacher use of behavior-specific praise. *Behavior Modification*, 31, 247–263.
- Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. In S. Hatzipanagos & R. Rochon (Eds.), *Approaches to assessment that enhance learning in higher education* (pp. 45–60). Routledge. <https://doi.org/10.1080/02602930903541015>

- Schleppenbach, M., Flevares, L. M., Sims, L. M., & Perry, M. (2007). Teachers' responses to student mistakes in Chinese and US mathematics classrooms. *The Elementary School Journal*, 108(2), 131-147.
- Saldaña, J. (2015). *The coding manual for qualitative researchers* (3rd ed.), Sage Publications Ltd.
- Scott, T. M., Gage, N., Hirn, R., & Han, H. (2019). Teacher and student race as a predictor for negative feedback during instruction. *School Psychology*, 34(1), 22.
- Senko, C. & Harackiewicz, J. M. (2005). Regulation of achievement goals: the role of competence feedback. *Journal of Educational Psychology*, 97(3), 320.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153-189.
- Skipper, Y., & Douglas, K. (2012). Is no praise good praise? Effects of positive feedback on children's and university students' responses to subsequent failures. *British Journal of Educational Psychology*, 82(2), 327-339.
- Smit, R., Bachmann, P., Dober, H., & Hess, K. (2024). Feedback levels and their interaction with the mathematical reasoning process. *The Curriculum Journal*, 35(2), 184-202.
- Starck, J. G., Riddle, T., Sinclair, S., & Warikoo, N. (2020). Teachers are people too: Examining the racial bias of teachers compared to other American adults. *Educational Researcher*, 49(4), 273-284.

- Stovner, R. B., & Klette, K. (2022). Teacher feedback on procedural skills, conceptual understanding, and mathematical practices: A video study in lower secondary mathematics classrooms. *Teaching and Teacher Education, 110*, 103593.
- Stroet, K., Opdenakker, M.-C. & Minnaert, A. (2013). Effects of need supportive teaching on early adolescents' motivation and engagement: A review of the literature. *Educational Research Review, 9*, 65–87.
- Sutherland, K. S., Wehby, J. H., & Copeland, S. R. (2000). Effect of varying rates of behavior-specific praise on the on-task behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 8*(1), 2-8.
- Thacker, I., Copur-Gencturk, Y., & Cimpian, J. R. (2022). Teacher bias: A discussion with special emphasis on gender and STEM learning. In T. L. Good & M. McCaslin (Eds.), *The Routledge Encyclopedia of Education: Educational Psychology Edition*. Routledge.
- Tiedemann, J. (2000). Gender-related beliefs of teachers in elementary school mathematics. *Educational Studies in Mathematics, 41*(2), 191-207.
- University of Connecticut. (n.d.). *Strategies to increase the utility value of your class*. Retrieved, September, 10, 2026. https://nrcgt.uconn.edu/underachievement_study/goal-valuation/gv_goalva04/
- Walton, G. M. & Crum, A. J. (2021). *Handbook of wise interventions*. The Guilford Press.
- Wisniewski, B., Zierer, K. & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*, 3087.

- Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, 47, 302–314.
- Yeager, D. S. Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., ... & Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology*, 143(2), 804–824.
- Yeager, D. S., Purdie-Vaughns, V., Hooper, S. Y. & Cohen, G. L. (2017). Loss of institutional trust among racial and ethnic minority adolescents: A consequence of procedural injustice and a cause of life-span outcomes. *Child Development*, 88, 658–676.
- Zentall, S. R., & Morris, B. J. (2010). “Good job, you’re so smart”: The effects of inconsistency of praise type on young children’s motivation. *Journal of Experimental Child Psychology*, 107(2), 155-163.

Table 1. *Background Characteristics of Teachers in the Present Sample Compared with a Nationwide Sample*

Teacher or school characteristic	Analytic sample (%) <i>N</i> = 963	Control group (%) <i>N</i> = 510	Treatment group (%) <i>N</i> = 453	Nationwide sample of U.S. public school teachers (%)
Sex ^a				
Female	82.9	81.6	84.3	76.8
Male	14.1	15.3	12.8	23.2
Prefer not to say	3.0	3.1	2.9	N/A
Race/ethnicity ^a				
White	68.9	66.7	71.3	79.9
Black	8.2	8.8	7.5	6.1
Hispanic	8.7	9.8	7.5	9.4
Other	14.2	14.7	13.7	4.6
Regular certification ^a	94.1	93.1	95.1	89.8
Years of teaching experience ^a				
Less than 3	5.6	4.3	7.1	7.3
3 to 9	33.8	34.5	32.9	29.1
10 to 20	37.3	39.0	35.3	37.3
More than 20	23.4	22.2	24.7	26.3
School region ^b				
Midwest	13.1	12.0	14.4	21.4
Northeast	11.7	12.2	11.3	19.5
South	38.9	41.6	36.0	40.3
West	36.2	34.3	38.4	18.8

Note. *N* = 963 teachers in the analytic sample. The Other race category included teachers from other races, such as Asian, Native American, or two or more races.

^aFrom *Digest of Education Statistics*, National Center for Education Statistics, https://nces.ed.gov/programs/digest/d22/tables/dt22_209.22.asp.

^bFrom *Common Core of Data: America's Public Schools*, National Center for Education Statistics, https://nces.ed.gov/ccd/tables/201920_summary_2.asp.

Table 2. *Coding Areas for Teachers' Written Feedback in Response to Students' Mathematics Problems*

Attention to the strengths or weaknesses of student solutions		
Categories	Description of criteria	Examples
Both strengths and weaknesses	The teacher states strengths and weaknesses in the student's work or strategy.	<i>"Wow, I can really see you are working hard to make sense of this challenging math problem. While your answer of changing the last number in that pattern makes sense for 3, 6, 9, 12, ... look closely to see that there is another number in that pattern (4). How does the number pattern grow between each term? Do you see a pattern in the growth?"</i>
Only strengths	The teacher states only the strength(s) in the student's work or strategy without any reference to the weaknesses in it.	<i>"Excellent observation and this will apply for all numbers. Is there any other approaches to this problem that you can think of?"</i>
Only weaknesses	The teacher states only the weakness(es) in the student's work without any reference to the strengths in it.	<i>"His work is not accurate for each number in the problem because 3×1.5 does not equal 4. Also the work he showed does not show multiplying by 1.5 correctly.</i>
Neither weaknesses nor strengths	The teacher does not comment on weakness(es) or strength(s) in the student's mathematical work. The teacher may instead probe the student's thinking or provide a general comment.	<i>"Does this strategy work every single time? Could you show how it does each time?"</i>
Focus Level: Processes, Procedures, or Student Attributes		
Categories	Description of criteria	Examples
Process and strategies	The teacher focuses on the processes of mathematical thinking and/or <i>how</i> to solve the problem.	<i>"You said multiply which I could understand because you noticed the numbers are going up. How did you arrive at 1.5?"</i>

Steps and calculations	The teacher focuses on <i>what</i> to do to solve the problem.	<i>"Yes, you figured out that the rule is to increase the amount you add by one each time to get the next number in the pattern."</i>
Student attributes	The teacher focuses on the student with positive or negative evaluations rather than on the student's work.	<i>"Although the student does not have the correct answer, she developed an explanation that was beyond the average ability of most 5th graders. Multiplying by 1.5 isn't a common choice a 5th grade student would make, and the math almost works out perfectly. Impressive."</i>
Generic	The teacher's response is generic and does not focus on a particular process, task-level step, or person.	<i>"Good job."</i>
Teacher's Comments on Student Ability and Effort		
Categories	Description of criteria	Examples
Positive Ability	The teacher indicated a positive statement of the students' mathematical ability and/or competence.	<i>"Although the student does not have the correct answer, she developed an explanation that was beyond the average ability of most 5th graders. Multiplying by 1.5 isn't a common choice a 5th grade student would make, and the math almost works out perfectly. Impressive."</i>
No/Mixed Ability	The teacher indicates either no positive/negative statement of the students' mathematical ability and/or competence <i>or</i> provides a statement that is both positive and negative.	<i>"Dear Student, Thank you for trying your best. I want you to try harder next time. Showing your work is your first step. Also, using the strategies I have taught you will help you solve the problems on the test. I have very high standards and I know you can meet my expectations. Keep practicing and you will get it. Tutoring will help."</i>

Negative Ability	The teacher indicated a negative statement of the students' mathematical ability and/or competence.	<i>"This student was extremely confused, as seen in his work and that he crossed it out. This student needs to be pulled for reteaching and small group instruction."</i>
Positive Effort	The teacher indicated a positive statement of the effort students put into solution.	<i>"I love it! You must have worked really hard on this."</i>
No/Mixed Effort	The teacher does not indicate a positive or negative statement of the effort students put into a solution <i>or</i> teacher indicates both positive and negative statements.	<i>Incredible observation! This rule would work for the first 4 numbers in the pattern, but 9×1.5 does not equal 13. See if you can find another way to explain this pattern! You're on the right track.</i>
Negative Effort	The teacher indicated a negative statement of the effort students put into solution.	<i>"You need to show more work with your problem! If you multiply by 1.5, the math would not equal up! Let's try this together by separating the numbers."</i>

Teacher's Analysis of the Student Work		
Categories	Description of criteria	Examples
Specific and correct analysis	The teacher explicitly and correctly analyzes the student's strategy with enough details about what is correct or incorrect in the student's work.	<i>"$3 \times 1.5 = 4.5$ and that breaks the sequence. 6×1.5 works but if you used your rule of $1.5x$, 6 would not appear in the sequence."</i>

Implicit and correct analysis	The teacher's response does not explicitly state what is incorrect or correct in the student's work, Analysis may include specific reference to the student's work or probing questions for the student to evaluate his or her work.	<i>"Great start, please provide more explanation. You are adding one, but to what?"</i>
Incorrect or general analysis	The teacher's response indicates that the teacher misunderstood the student's solution or that the teacher made no specific comment on the student's work.	<i>"Your answer is correct. You can multiply each number by 1.5 to complete the series; however, your explanation is missing specific details. How did you figure out the pattern?" (the student work is incorrect.)</i>
Specificity of Suggestions to Improve the Student's Understanding		
Categories	Description of criteria	Examples
Specific limitation and targeted suggestion	The teacher provides targeted suggestions to improve the student's understanding, with explicit connections to the limitations in the student's understanding.	<i>"Remember, a pattern rule should fit with the entire pattern as you move from term to term. If the rule you wrote was applied to 3, it would not result in the next term 4."</i>
Specific limitation addressed with no or vague suggestion	The teacher points out specific limitations in the student work, but (a) the teacher does not provide any suggestions to improve it or (b) the provided suggestion is vaguely connected to the key issues the teacher identified.	<i>"I see that you think we should multiply each number by 1.5. Why don't you do that with 3 to show me what your result would be. If we multiply 3 by 1.5, our answer would be 4.5. Does that fit the given pattern?"</i>
Targeted suggestion but no or vague connection to the student's work	The teacher provides targeted suggestions to overcome the student's misconception, but the student's misconception is not pointed out; no explicit connection is made to the student's work.	<i>"Have the student explain the rule from the beginning—How do you go from 3 to 4? Then from 4 to 6? Then from 6 to 9? Each time you add, what are you adding by?"</i>

Generic or no suggestions

The teacher does not provide any targeted suggestions to improve the student's existing understanding and does not talk about any specific limitations in the student's work.

"You did a great job identifying the rule in the growing pattern. I like that you used both numbers and words to explain the rule. You even gave an example to support your writing!"

Incorrect suggestion

The teacher provides incorrect suggestions or guidance to improve the student's existing understanding.

"Your answer is correct. You are adding one each time. You have figured out the pattern. Good job!"

Table 3. *Adjusted Average Proportion of Teacher Feedback Categories by Treatment and Control Conditions*

Category	Control group		Treatment group		RR
	N = 510		N = 453		
	Mean	SE	Mean	SE	
Attention to the strengths or weaknesses of student solutions					
Both	0.264	0.012	0.386	0.014	1.46 *
Only strengths	0.176	0.009	0.171	0.010	0.97
Only weaknesses	0.400	0.011	0.358	0.012	0.90
Neither	0.160	0.010	0.086	0.008	0.54 *
Overall emphasis					
Processes	0.149	0.010	0.143	0.010	0.96
Steps	0.723	0.011	0.710	0.013	0.98
Students	0.006	0.002	0.042	0.007	7.00 *
Generic	0.122	0.010	0.106	0.010	0.87
Comments on student ability or effort					
Positive ability	0.021	0.003	0.081	0.009	3.86 *
No/mixed ability	0.974	0.004	0.893	0.010	0.92 *
Negative ability	0.005	0.002	0.027	0.005	5.40 *
Positive effort	0.022	0.004	0.035	0.005	1.59
No/mixed effort	0.973	0.005	0.937	0.007	0.96 *
Negative effort	0.005	0.002	0.028	0.004	5.60 *
Teacher's analysis of the student work					
Explicit	0.278	0.012	0.313	0.013	1.13
Implicit	0.499	0.012	0.476	0.013	0.95
Generic/incorrect	0.224	0.011	0.211	0.013	0.94
Suggestions to improve student learning					
Targeted suggestions for the identified area	0.108	0.008	0.129	0.009	1.19
Only the area for improvement	0.159	0.008	0.153	0.009	0.96
Only targeted suggestions	0.189	0.010	0.182	0.010	0.96
Generic/no suggestions	0.529	0.013	0.518	0.014	0.98
Incorrect suggestions	0.015	0.003	0.018	0.003	1.20

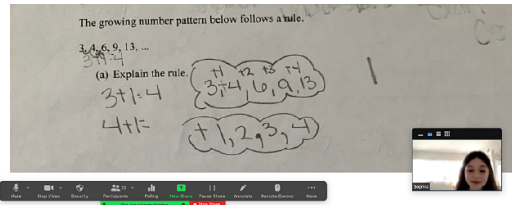
Note. RR is the relative ratio of the outcome in the treatment condition vs. the control condition. An asterisk (*) indicates the RR is significant after the Bonferroni correction.

Figure 1.

Sample Screenshots from the Control and Treatment Conditions for Black and White students.

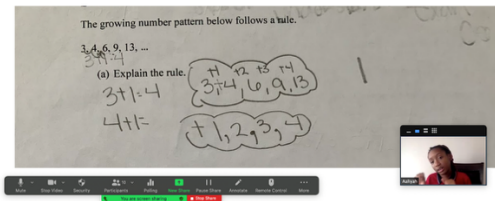
Control Condition Prompt

Black
Student
image



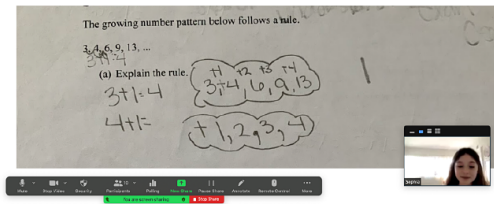
Provide specific feedback to this student in the way that you normally would.

White
Student
image

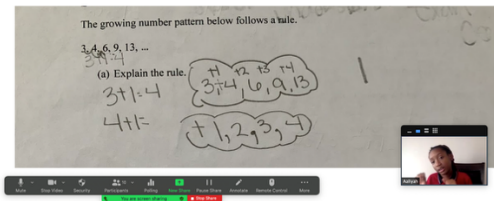


Provide specific feedback to this student in the way that you normally would.

Treatment Condition Prompt



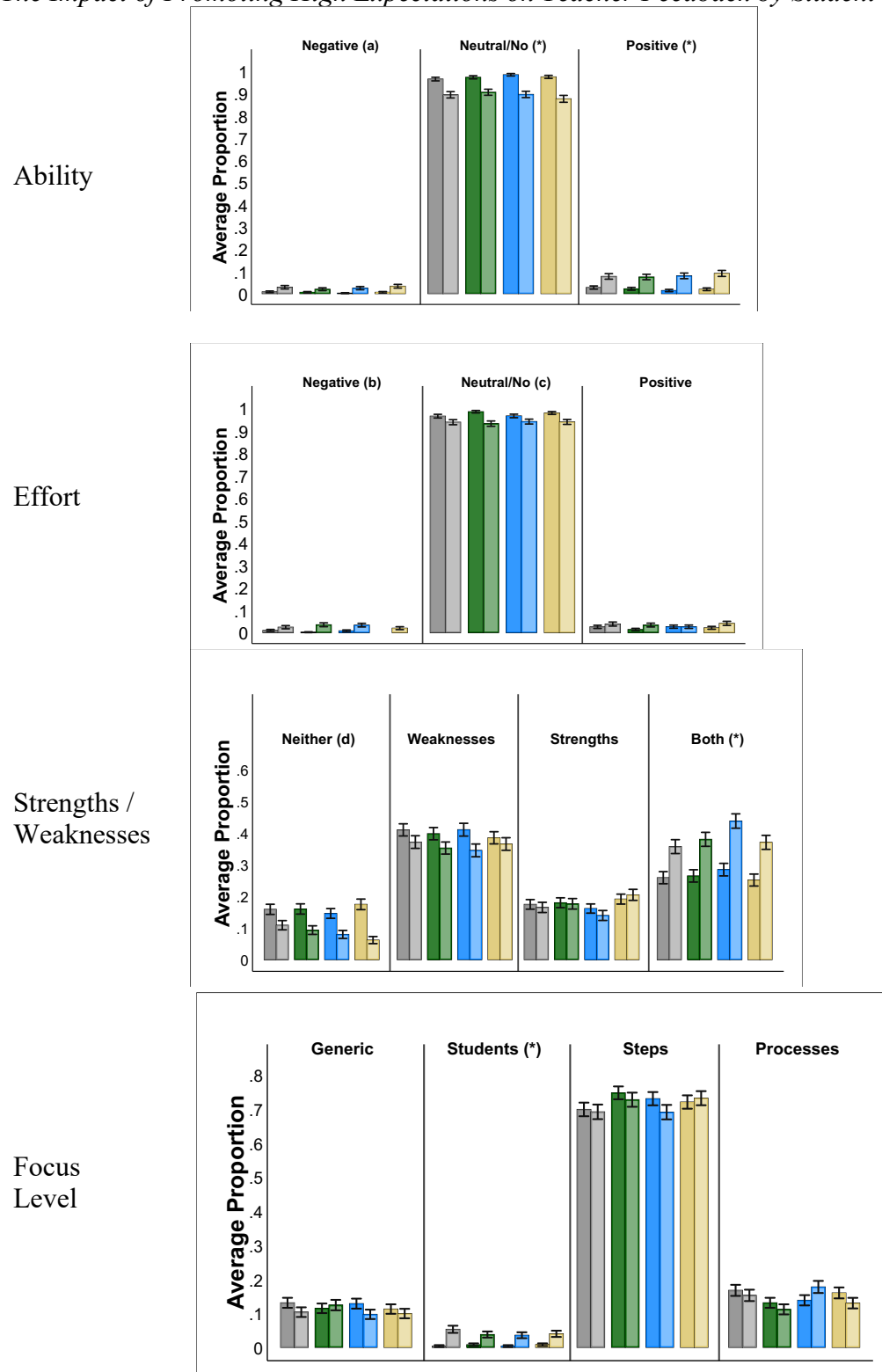
Provide specific feedback to this student, emphasizing your high standards and that the student has the potential to meet those standards.



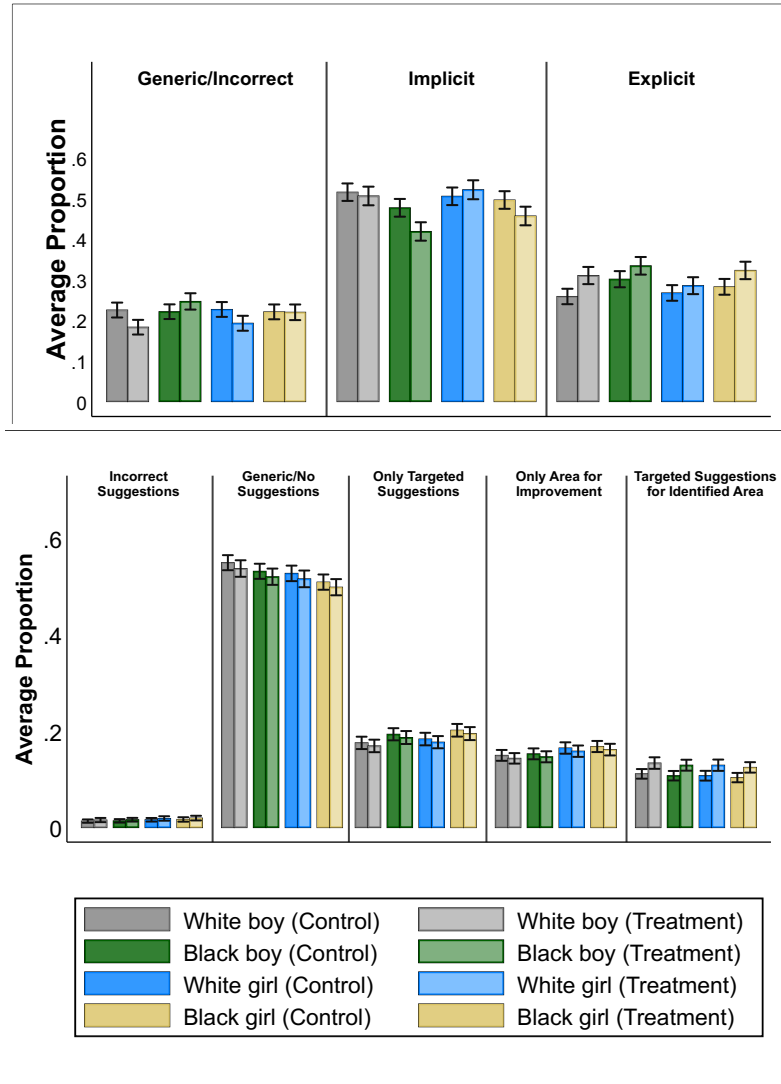
Provide specific feedback to this student, emphasizing your high standards and that the student has the potential to meet those standards.

Figure 2.

The Impact of Promoting High Expectations on Teacher Feedback by Student Gender and Race.



Analysis of Student Solutions



Note. The error bars are +/- 1 standard error. An asterisk (*) indicates the differences are significant after the Bonferroni correction. a = The differences are significant for White and Black girls after the Bonferroni correction. b = The differences are significant for all groups except White boys ($p = .08$). c = The differences are significant for Black girls and Black boys after the Bonferroni correction. d = The differences are significant for all groups except White boys after the Bonferroni correction ($p = .02$).

Online Supplementary Materials

Supplementary Text 1. Rubric for Coding the Subject of the Action

Code 1: No Third-Person Feedback

The response is written directly to the student or framed from the teacher perspective in a way that addresses or engages the student's work.

This can include responses that use:

- Second-person language (you, your)
- First-person or language suggesting both teacher/student (e.g., I, we, us) or "let's" when they are referencing working together.
- Direct prompts or questions (e.g., try, look, check, explain)

Examples:

- "I see you have an understanding by looking at your numbers. Tell me more about 'adding one every time.'"
- "Excellent job! You correctly saw the pattern and explained your thinking clearly and accurately - keep up the great work!"
- "Next time I want you to show me your work, and show me how you figured out it would be multiply times 3. We need to also remember we cannot change the problem to match what we want, so how can you change your thinking using the numbers that are given to you?"
- Let's look at all of our numbers again. 3, 4, 6, 9, 13. While there are several numbers that include factors of 3, let's take a closer look at the number 4. Does the rule work for this number? The rule also did not apply to 13. Rules must work for each number in our pattern. What are some other strategies we can use to help us identify a rule that works for all of the numbers listed?

Code 2: Contain a Mix of Person-Centered Feedback

Feedback can include setup language such as "I would ask..." and then include feedback language or questions that the teacher would use with the student.

Examples:

- "As her teacher, I would have asked her to use words to explain her thinking. She did a great job showing it with numbers, but I would have told her mathematicians can also use words to explain this. For example, you need to add one more that the number you added before to the next number."
- "I would ask the student to show me their work and try it out. Let's multiply $3 * 1.5$. What do we get? So does your rule work? Why not?"

- “I would ask this student to be more specific in their explanation--"To what are you adding 1? If they weren't sure, I would supply the words increase and addend to help this child improve their explanation.”
- “I would ask the child to explain his answer. You got the answer correct, but can you explain how you got this. Most time, students are able to explain the patterns. Time after time, I see students try to incorporate things that they learned, example this student uses an array 9×4 and division. Even though 9×4 and $36/2$ are no where in the problem. I think it's okay because this child maybe expressing their thought about something they learned previously. The "now what" on the paper made me think that the student may or may not know what he talking about. I would ask this student to explain. Sometimes, just asking them out loud can get a totally different thought from the student. Each child is different.”

Code 3: Third Person

Responses are written about the student rather than directed to the student. They may describe the student's thinking, performance, misunderstanding, or work from an observational perspective. This includes feedback that uses language such as “the student, the child, he/him, she/her” or references the student as an object of discussion and does not use any language that suggests the actual student is the intended audience for the feedback.

Responses should be coded as 3 when they are primarily descriptive or diagnostic rather than written as feedback that could be directly given to the student. If the feedback identifies the student's issue or describes a general teacher action but does not include a clear student-facing prompt, question, explanation, or next step, it should be coded as 3.

Examples:

- “The student may not understand what a pattern is or how it relates to the sequence of numbers.”
- “Couldn't follow the strategy completely. She seemed to arrive at 12 and 24, but sometimes used the = incorrectly.”
- “This student again seemed to understand the solution but then made an error which showed she was not proficient in her answers. I would pull this student in a small group setting to reteach this lesson.”
- “He clearly has an understanding of addition and subtraction principles but is confused where to go next. He needs help in applying multiplication and division principles. By showing the array for 9×6 , you can tell he has basic understanding but does not know how to apply this to the problem.”
- “This student is correct but they need to explicitly say that they are adding one to the previous pattern each time.”
- “They did a great job answering the problem.”

Online Supplementary Tables

Table S1. Multinomial Logistic Regression Results for Teachers' Attention to the Strengths or Weaknesses of Student Solutions

<i>Predictor</i>	Category								
	Neither			Only Strengths			Both Strengths and Weaknesses		
	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c
Treatment	-0.95*** (0.25)	-1.10*** (0.28)	-1.17*** (0.28)	0.22 (0.21)	0.20 (0.23)	0.22 (0.23)	0.50** (0.17)	0.52** (0.18)	0.48** (0.18)
Boys	-0.13 (0.17)	-0.16 (0.17)	-0.19 (0.18)	-0.12 (0.17)	-0.16 (0.18)	-0.18 (0.18)	0.00 (0.15)	-0.04 (0.16)	-0.08 (0.16)
Treatment × Boys	0.57~ (0.29)	0.58~ (0.32)	0.64~ (0.33)	-0.02 (0.25)	-0.02 (0.27)	-0.04 (0.27)	0.04 (0.20)	0.06 (0.21)	0.13 (0.21)
White	-0.27 (0.17)	-0.23 (0.18)	-0.21 (0.18)	-0.30 (0.19)	-0.21 (0.20)	-0.23 (0.20)	0.02 (0.16)	0.04 (0.17)	-0.00 (0.17)
Treatment × White	0.56~ (0.32)	0.62~ (0.36)	0.54 (0.37)	-0.09 (0.30)	-0.27 (0.32)	-0.31 (0.32)	0.17 (0.22)	0.12 (0.23)	0.18 (0.24)
White × Boys	0.22 (0.22)	0.22 (0.23)	0.27 (0.23)	0.00 (0.21)	0.00 (0.21)	0.23 (0.25)	-0.09 (0.21)	-0.07 (0.21)	-0.03 (0.22)
Treatment × White × Boys	-0.43 (0.40)	-0.65 (0.43)	-0.55 (0.44)	(0.23) 0.00	(0.24) -0.08	-0.03 (0.38)	-0.25 (0.28)	-0.29 (0.29)	-0.32 (0.29)
Intercept	-1.46*** (0.38)	-1.72*** (0.41)	-1.60*** (0.41)	-0.74* (0.32)	-0.94** (0.34)	-0.87* (0.34)	0.03 (0.29)	-0.08 (0.31)	0.02 (0.31)

Note. The reference category is Only Weaknesses. Values in parentheses indicate robust standard errors clustered at the teacher level. The model also includes the items as fixed effects. *Third-person feedback removed* refers to excluding responses written entirely in the third person. Responses containing a combination of third-person statements (e.g., descriptions of planned teacher actions) and direct feedback addressed to the student were retained. *Only direct feedback included* refers to retaining only teacher feedback addressed directly to the student; therefore, teacher feedback containing teacher-directed statements or descriptions of planned teacher actions was excluded.

^a $N = 963$, ^b $N = 894$, ^c $N = 886$ teachers; ^a $N = 3852$, ^b $N = 3471$, ^c $N = 3426$ feedback responses across all teachers.

~ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$. The significance levels were not adjusted for multiple comparisons.

Table S2. Multinomial Logistic Regression Results for Teachers' Overall Emphasis

Predictor	Category								
	Generic			Students			Processes		
	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c
Treatment	-0.14 (0.22)	-0.20 (0.23)	-0.24 (0.23)	1.60** (0.56)	1.55 (1.13)	1.23 (1.16)	-0.23 (0.19)	-0.19 (0.20)	-0.21 (0.20)
Boys	-0.02 (0.18)	-0.01 (0.18)	-0.02 (0.18)	-0.05 (0.62)	0.65 (1.23)	0.65 (1.23)	-0.24 (0.16)	-0.24 (0.17)	-0.22 (0.17)
Treatment × Boys	0.25 (0.26)	0.15 (0.29)	0.14 (0.28)	-0.00 (0.66)	-0.37 (1.29)	-0.07 (1.29)	0.10 (0.24)	0.14 (0.25)	0.17 (0.25)
White	0.12 (0.17)	0.15 (0.18)	0.12 (0.18)	-0.73 (0.87)	-0.06 (1.41)	-0.07 (1.41)	-0.16 (0.17)	-0.09 (0.17)	-0.10 (0.17)
Treatment × White	-0.08 (0.26)	-0.08 (0.28)	-0.08 (0.28)	0.69 (0.92)	0.73 (1.50)	0.69 (1.55)	0.54* (0.24)	0.44~ (0.25)	0.46~ (0.26)
White × Boys	0.09 (0.24)	0.06 (0.25)	0.07 (0.24)	0.10 (1.18)	0.13 (1.74)	0.13 (1.74)	0.48* (0.22)	0.42~ (0.22)	0.36 (0.23)
Treatment × White × Boys	-0.25 (0.34)	-0.41 (0.37)	-0.36 (0.37)	0.35 (1.22)	-0.44 (1.80)	-0.54 (1.80)	-0.49 (0.33)	-0.42 (0.34)	-0.38 (0.34)
Intercept	-1.79*** (0.30)	-1.86*** (0.33)	-1.82*** (0.32)	-6.01*** (0.84)	-7.94*** (1.59)	-7.54*** (1.60)	-2.56*** (0.30)	-2.59*** (0.31)	-2.66*** (0.32)

Note. The reference category is Steps. Values in parentheses indicate robust standard errors clustered at the teacher level. The model also includes the items as fixed effects. *Third-person feedback removed* refers to excluding responses written entirely in the third person. Responses containing a combination of third-person statements (e.g., descriptions of planned teacher actions) and direct feedback addressed to the student were retained. *Only direct feedback included* refers to retaining only teacher feedback addressed directly to the student; therefore, teacher feedback containing teacher-directed statements or descriptions of planned teacher actions was excluded.

^a $N = 963$, ^b $N = 894$, ^c $N = 886$ teachers; ^a $N = 3852$, ^b $N = 3471$, ^c $N = 3426$ feedback responses across all teachers.

~ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$. The significance levels were not adjusted for multiple comparisons.

Table S3. Multinomial Logistic Regression Results for Teachers' Comments on Student Ability or Effort

Predictor	Category											
	Negative ability			Positive ability			Negative Effort			Positive effort		
	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c
Treatment	1.86** (0.63)	14.88*** (0.53)	14.61*** (0.61)	1.65*** (0.36)	1.37*** (0.37)	1.30*** (0.38)	14.77*** (0.37)	14.98*** (0.33)	14.97*** (0.41)	0.71~ (0.39)	0.81* (0.39)	0.79* (0.39)
Boys	-0.01 (0.67)	13.88*** (0.73)	13.22*** (1.03)	0.10 (0.43)	-0.10 (0.44)	-0.11 (0.45)	12.37*** (1.01)	12.61*** (0.99)	12.61*** (1.02)	-0.44 (0.44)	-0.45 (0.44)	-0.46 (0.44)
Treatment × Boys	-0.55 (0.76)	14.61*** (1.12)	-13.65*** (1.36)	-0.33 (0.46)	0.08 (0.48)	0.10 (0.48)	-11.79*** (1.08)	-12.30*** (1.07)	-12.21*** (1.09)	0.24 (0.52)	0.12 (0.52)	0.13 (0.52)
White	-1.09 (1.16)	13.20*** (1.02)	13.24*** (1.03)	-0.36 (0.50)	-0.34 (0.50)	-0.35 (0.50)	13.83*** (0.52)	14.04*** (0.49)	14.05*** (0.54)	0.23 (0.34)	0.24 (0.34)	0.22 (0.34)
Treatment × White	0.76 (1.21)	12.98*** (1.15)	-13.25*** (1.15)	0.20 (0.55)	0.27 (0.56)	0.24 (0.56)	-13.30*** (0.65)	-13.62*** (0.64)	-13.63*** (0.68)	-0.69 (0.51)	-0.77 (0.51)	-0.75 (0.51)
White × Boys	1.40 (1.31)	13.19*** (1.44)	-12.55*** (1.62)	0.61 (0.58)	0.80 (0.58)	0.81 (0.58)	-12.16*** (1.22)	-12.39*** (1.20)	-12.39*** (1.23)	0.41 (0.49)	0.34 (0.49)	0.27 (0.51)
Treatment × White × Boys	-0.69 (1.38)	13.40*** (1.80)	12.98*** (1.93)	-0.40 (0.63)	-0.83 (0.65)	-0.81 (0.65)	11.25*** (1.34)	11.54*** (1.37)	11.28*** (1.40)	0.17 (0.64)	0.32 (0.65)	0.37 (0.67)
Intercept	-6.08*** (1.06)	20.72*** (1.43)	-20.49*** (1.39)	-4.83*** (0.56)	-4.57*** (0.61)	-4.42*** (0.61)	-18.94*** (0.96)	-19.24*** (1.01)	-18.91*** (0.97)	-5.40*** (0.86)	-5.69*** (0.92)	-5.65*** (0.92)

Note. The reference category is No/Mixed. Values in parentheses indicate robust standard errors clustered at the teacher level. The model also includes the items as fixed effects. *Third-person feedback removed* refers to excluding responses written entirely in the third person. Responses containing a combination of third-person statements (e.g., descriptions of planned teacher actions) and direct feedback addressed to the student were retained. *Only direct feedback included* refers to retaining only teacher feedback addressed directly to the student; therefore, teacher feedback containing teacher-directed statements or descriptions of planned teacher actions was excluded.

^a $N = 963$, ^b $N = 894$, ^c $N = 886$ teachers; ^a $N = 3852$, ^b $N = 3471$, ^c $N = 3426$ feedback responses across all teachers.

~ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$. The significance levels were not adjusted for multiple comparisons.

Table S4. *Multinomial Logistic Regression Results for Teachers' Analysis of the Student Work*

<i>Predictor</i>	Category					
	Generic/Incorrect			Explicit		
	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c
Treatment	0.08 (0.17)	-0.08 (0.18)	-0.09 (0.18)	0.22 (0.15)	0.17 (0.16)	0.14 (0.16)
Boys	0.04 (0.14)	0.06 (0.15)	0.04 (0.15)	0.11 (0.13)	0.10 (0.14)	0.08 (0.14)
Treatment × Boys	0.16 (0.21)	0.22 (0.23)	0.23 (0.23)	0.02 (0.20)	0.07 (0.21)	0.12 (0.21)
White	0.01 (0.15)	0.02 (0.16)	0.00 (0.16)	-0.08 (0.14)	-0.06 (0.15)	-0.07 (0.15)
Treatment × White	-0.28 (0.22)	-0.20 (0.25)	-0.24 (0.25)	-0.19 (0.21)	-0.17 (0.22)	-0.16 (0.22)
White × Boys	-0.07 (0.19)	-0.07 (0.20)	-0.05 (0.20)	-0.16 (0.20)	-0.15 (0.20)	-0.19 (0.20)
Treatment × White × Boys	-0.16 (0.29)	-0.28 (0.33)	-0.27 (0.33)	0.15 (0.28)	0.07 (0.30)	0.08 (0.30)
Intercept	-1.35*** (0.26)	-1.35*** (0.29)	-1.32*** (0.29)	-1.75*** (0.26)	-1.64*** (0.27)	-1.61*** (0.27)

Note. The reference category is Implicit. Values in parentheses indicate robust standard errors clustered at the teacher level. The model also includes the items as fixed effects. *Third-person feedback removed* refers to excluding responses written entirely in the third person. Responses containing a combination of third-person statements (e.g., descriptions of planned teacher actions) and direct feedback addressed to the student were retained. *Only direct feedback included* refers to retaining only teacher feedback addressed directly to the student; therefore, teacher feedback containing teacher-directed statements or descriptions of planned teacher actions was excluded.

^a $N = 963$, ^b $N = 894$, ^c $N = 886$ teachers; ^a $N = 3852$, ^b $N = 3471$, ^c $N = 3426$ feedback responses across all teachers.

$\sim p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$. The significance levels were not adjusted for multiple comparisons.

Table S5. *Multinomial Logistic Regression Results for Teachers' Suggestions to Improve Student Learning*

Predictor	Category											
	Incorrect Suggestions			Only Targeted Suggestions			Only the Area for Improvement			Targeted Suggestions for the Identified Area		
	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c	All feedback ^a	Third-person feedback removed ^b	Only direct feedback included ^c
Treatment	-0.12 (0.45)	-0.33 (0.51)	-0.36 (0.51)	0.14 (0.18)	0.32~ (0.18)	0.36~ (0.19)	0.17 (0.20)	0.27 (0.21)	0.24 (0.21)	0.15 (0.22)	0.31 (0.23)	0.31 (0.23)
Boys	-0.68 (0.48)	-0.69 (0.48)	-0.71 (0.48)	0.02 (0.15)	-0.00 (0.16)	0.05 (0.16)	-0.01 (0.18)	-0.03 (0.19)	-0.07 (0.19)	-0.01 (0.20)	-0.03 (0.20)	-0.03 (0.20)
Treatment × Boys	0.05 (0.75)	0.24 (0.82)	0.26 (0.82)	-0.18 (0.22)	-0.16 (0.23)	-0.20 (0.23)	-0.24 (0.26)	-0.25 (0.28)	-0.19 (0.28)	-0.00 (0.28)	-0.02 (0.29)	-0.01 (0.30)
White	-0.73 (0.49)	-0.90~ (0.52)	-0.91~ (0.52)	0.02 (0.16)	-0.01 (0.17)	0.03 (0.17)	0.07 (0.18)	0.04 (0.19)	0.03 (0.19)	-0.05 (0.20)	-0.08 (0.20)	-0.10 (0.21)
Treatment × White	0.56 (0.68)	1.01 (0.74)	1.04 (0.74)	-0.28 (0.25)	-0.20 (0.26)	-0.22 (0.26)	-0.21 (0.27)	-0.22 (0.29)	-0.16 (0.29)	0.12 (0.29)	0.14 (0.30)	0.21 (0.31)
White × Boys	0.82 (0.73)	1.01 (0.74)	1.01 (0.74)	-0.20 (0.21)	-0.22 (0.22)	-0.27 (0.23)	-0.15 (0.25)	-0.16 (0.26)	-0.17 (0.26)	-0.07 (0.27)	-0.06 (0.27)	-0.06 (0.29)
Treatment × White × Boys	0.01 (1.04)	-0.40 (1.10)	-0.45 (1.10)	0.33 (0.31)	0.35 (0.33)	0.36 (0.33)	0.16 (0.37)	0.30 (0.39)	0.20 (0.40)	0.06 (0.39)	0.17 (0.40)	0.12 (0.41)
Intercept	-2.86*** (0.87)	-2.42** (0.89)	-2.42** (0.89)	1.23*** (0.32)	1.35*** (0.34)	1.20*** (0.34)	-1.87*** (0.33)	-1.80*** (0.35)	-1.76*** (0.35)	0.44 (0.35)	0.55 (0.36)	0.62 (0.38)

Note. The reference category is Generic/No suggestions. Values in parentheses indicate robust standard errors clustered at the teacher level. The model also includes the items as fixed effects. *Third-person feedback removed* refers to excluding responses written entirely in the third person. Responses containing a combination of third-person statements (e.g., descriptions of planned teacher actions) and direct feedback addressed to the student were retained. *Only direct feedback included* refers to retaining only teacher feedback addressed directly to the student; therefore, teacher feedback containing teacher-directed statements or descriptions of planned teacher actions was excluded.

^a $N = 963$, ^b $N = 894$, ^c $N = 886$ teachers; ^a $N = 3852$, ^b $N = 3471$, ^c $N = 3426$ feedback responses across all teachers.

~ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$. The significance levels were not adjusted for multiple comparisons.