Are Students' Math and Verbal Motivational Beliefs Malleable? The Role of Praise in Dimensional Comparisons

Kexin Ren[1], Amanda Grenell[2], & Elizabeth A. Gunderson[2]
Temple University[1]
Indiana University[2]
fsu

Word count (main text + references): 10,523

Corresponding author information:

Elizabeth Gunderson
Department of Psychological and Brain Sciences, Indiana University
1101 E. 10th Street
Bloomington, IN 47405
Telephone: (812) 855-5183
Electronic mail: eagunder@iu.edu

Authors note:

This study was part of the doctoral dissertation completed by Kexin Ren. This study was not preregistered. All data, analysis code, and research materials are publicly available at https://osf.io/bpq49/?view_only=0dbdfbb1f7684840a384b4252e8dd3ec.
This research was supported by NSF ECR-1760144 to Elizabeth Gunderson. We thank Annemarie Hindman, Kathy Hirsh-Pasek, Vishnu Murty, Nora Newcombe, and Marsha Weinraub for their helpful feedback on an earlier version of this manuscript. We also thank the students and families who gave their time to this research and the research assistants who helped collect the data: Victoria Bartek, Kimberly Bohl, Melissa Mildort, and Jorge Carvalho Pereira.

Abstract

To determine their academic strengths and weaknesses, students compare their own performance across domains (e.g., math versus English), a process referred to as *dimensional comparisons*. For example, individuals' higher-scoring English performance may negatively affect their math motivational beliefs (competence self-concepts and intrinsic values), resulting in favoritism toward English. Students' motivation can also be affected by praise from adults. However, praise in one domain (e.g., English) may have unexpected negative effects on motivation in the contrasting domain (e.g., math) through dimensional comparisons. We experimentally investigated the impact of receiving praise in only one domain on students' domain-specific motivational beliefs. We hypothesized that students would have higher motivational beliefs in the praised domain and lower motivational beliefs in the non-praised domain compared to students who received no praise. Seventh-to-9th-graders (10-15-year-olds; N=108; 46 girls; 92 living in the U.S.; 84.8% White, 2.9% Asian or Asian American, 2.9% Black or African American, 9.5% multiple races; parents' education range: 13-18 years) showed heightened verbal competence self-concepts after receiving praise on either verbal or math performance. College students (first-to-5th-year; N=109; 89 females; 105 living in the U.S.; 58.9% White, 21.5% Asian or Asian American, 10.3% Black or African American, 5.6% multiple races, 3.7% other races) showed higher verbal intrinsic values after receiving praise on verbal performance. Results supported positive effects of praise in the verbal domain only and were inconsistent with the predicted negative effects on the non-praised domain. We suggest that students' verbal motivational beliefs are more malleable than math beliefs when receiving disproportionate praise.

Keywords: dimensional comparisons, praise, competence self-concepts, intrinsic values

Are Students' Math and Verbal Motivational Beliefs Malleable? The Role of Praise in Dimensional Comparisons

Understanding the development of academic motivation is critical for encouraging students' achievement and opportunities. Academic motivation is known to be domain-specific, can differ across math and verbal domains, and is often lower in the math domain, especially for women and underrepresented minorities (e.g., Cvencek et al., in press; Nagy et al., 2006; Seo et al., 2019). The praise students receive for their academic performance is thought to have a major impact on students' motivation. Yet, research on the effects of praise on motivation has largely ignored the potential for praise to impact students' domain-specific motivational beliefs differently in the praised domain and in a non-praised domain. A growing literature on dimensional comparisons suggests that students contrast their own performance across domains (for a review, see Möller & Marsh, 2013), yet much of this research rests on correlational designs, leaving open the question of whether increasing students' motivational beliefs in one domain *causes* a decrease in a contrasting domain. Bringing these lines of research together, we experimentally tested whether praising students' performance in one domain would lead to an increase in motivational beliefs in that domain and a decrease in the contrasting domain, consistent with dimensional comparison theory. Understanding the interaction between praise and students' domain-specific motivation is critical for providing parents and teachers insights into the best ways to motivate students without potentially undermining students' motivation in other domains.

Dimensional Comparisons and Motivational Beliefs

People tend to choose a field of study in which they believe they possess the most skills, and that they place the highest interest and value in (Musu-Gillette et al., 2015). This process often involves comparing one's own performance across different domains, such as math and English. These intraindividual comparisons between domains are known as dimensional comparisons (Möller & Marsh, 2013). Through dimensional comparisons, individuals may shape their motivational beliefs – including competence self-concepts (beliefs about one's competence) and intrinsic values (enjoyment) – in one domain based on their performance in a different domain (Eccles, 2009; Möller & Köller, 2001; Möller & Marsh, 2013). In fact, there is evidence for these types of influences of dimensional comparisons starting in middle adolescence (Goetz et al., 2008; Marsh & Hau, 2004).For instance, in 15-year-olds from 26 countries, Marsh and Hau (2004) found that both math and verbal achievement were positively related to competence self-concepts in the same domain, but negatively related to the opposite domain. The effects of dimensional comparisons are found not only in competence self-concepts, but also in students' intrinsic values. Goetz et al. (2008) followed German students from grades 5 to 10 and found that math and language performance in the previous academic year positively predicted students' intrinsic values in the same class, but negatively predicted students' intrinsic values in contrasting-domain classes. Such comparison processes may result in favoritism towards the better-off domain compared to the worse-off domain. This can result in counterintuitive effects later in development, for example, leading people to opt out of science, technology, engineering, and mathematics (STEM) majors in college, because, despite being high achieving in math, they believe they are relatively better at non-STEM subjects such as English.

Disproportionate Praise and Dimensional Comparisons

In addition to internal dimensional comparison processes, environmental factors have also been shown to affect students' formation of motivational beliefs (Bong & Skaalvik, 2003; Gniewosz et al., 2014). A large amount of prior research has documented that adults transmit their beliefs about academic domains – including the importance of those domains and their beliefs about the child's ability in each domain – to children, either directly or indirectly through their behaviors (Gladstone et al., 2018; Gunderson et al., 2012; Pesu, Aunola, et al., 2016; Pesu et al., 2018; Pesu, Viljaranta, et al., 2016; Simpkins et al., 2012, 2015; Tiedemann, 2000). Feedback, as one of the important vehicles of adult-to-child transmission, has proven its powerful impacts (both positive and negative) on motivational beliefs (for a review, see Hattie & Timperley, 2007).

Several prior studies of dimensional comparisons have used experimental designs to demonstrate the causal effects of performance feedback on students' motivational beliefs (e.g., Möller & Köller, 2001; Pohlmann & Möller, 2009; Strickhouser & Zell, 2015). For example, Möller and Köller (2001) examined college students' math and verbal competence self-concepts after receiving performance feedback on math- and verbal-related tasks. Across three different experimental manipulations, the authors found consistent evidence that positive feedback (e.g., being told they scored above average) in math resulted in lower verbal competence self-concepts, with and without reference to the verbal domain.

These studies demonstrate that feedback that takes the form of a quantitative reflection of students' performance (e.g., "You ranked in the 83rd percentile of all the students" or "Your performance is above average") affected students' motivation in each domain via dimensional comparisons. However, in daily educational contexts, students receive not only quantitative feedback but also qualitative, subjective feedback from adults, such as praise (e.g., "You did a good job"). For example, Hyland and Hyland (2001) noted that almost half of the feedback teachers gave to college students was praise. Floress et al. (2018) also found that, across kindergarten to fifth grade, teachers on average provided 34.2 total praise statements (e.g., "Good job") per hour. Despite the ubiquity of praise in educational settings, it is unclear whether and how praise plays a role in dimensional comparisons.

Indeed, despite the large literature on praise, most research has been conducted within domains – either domain general (e.g., general intelligence) or domain specific (e.g., only in mathematics). Very few studies have examined the effects of praise across domains, even though adults have been shown to give disproportionate praise to students between academic domains and non-academic domains (Dweck et al., 1978). For example, in one study, girls received more non-academic praise (e.g., praising for neatness, behavior, and speaking clearly) than boys, which led girls to devalue the academic praise they received more than boys (Dweck et al., 1978). The potential negative effect of disproportionate praise raises the question of whether receiving disproportionate praise between academic domains (e.g., math versus English) would also undermine students' domain-specific motivational beliefs via dimensional comparisons.

Theoretically, receiving praise may trigger dimensional comparisons in several possible ways. To illustrate this, consider a student who has the same teacher for both math and English and who receives similar grades in both domains. Now imagine that the teacher praises the student's math performance but neglects to mention the student's English performance. There are at least three distinct inferences the student may make about the reason for the teacher's disproportionate praise. First, this student may infer that the teacher praised their math achievement because the teacher values math more than English. The student may then model the teacher's values and begin to value math more than English themself. This possibility is supported by research on expectancy-value theory (Eccles, 2009, 2015; Eccles et al., 1983), which posits that socializers (e.g., parents and teachers) play an important role in the formation of students' motivational beliefs, and that one way they can communicate their beliefs is by providing students with positive feedback.

A second possibility is that, after receiving disproportionate praise, this student may infer that the teacher believes that the student is higher-performing in math than in English, or is higher-performing in math than most of the other students in the same class. Thus, the teachers' praise may be interpreted as additional information, over and above the student's grade, indicating the students' high performance in math. This interpretation may lead the student to dimensionally compare their math and English performance and develop more positive motivational beliefs in math than in English.

A third possibility is that this student may infer that the teacher's disproportionate praise reflects the teacher's low expectations about their math abilities, which the student may interpret as the teacher praising the effort they have put in to perform well on math tasks. In contrast to the first two interpretations, this interpretation may result in *lower* motivational beliefs in math and potentially higher motivational beliefs in English. This possibility has been supported by some prior work arguing that praise for effort could suggest that individuals have to work hard to achieve because they lack abilities (for a review, see Amemiya & Wang, 2018). Praise for effort is most demotivating when students believe in an inverse relation between effort and ability (that high effort implies low ability and vice versa; Lam et al., 2008). This belief in the effort-ability relation varies across individuals and age groups, with children believing in a positive effort-ability relation in early elementary school, and a negative effort-ability relation by late elementary school (Barker & Graham, 1987; Muenks & Miele, 2017).

It is worth noting that the inferences the student makes are not necessarily aligned with the teacher's intentions. The first two inferences are likely to result in the same dimensional comparison patterns, where the praised domain is treated as the better-off domain. In contrast, the last inference is likely to lead to an opposite dimensional comparison pattern, where the praised domain is perceived as the worse-off domain.

An empirical study has provided some insights into how students in adolescence and early adulthood interpreted disproportionate praise (Möller, 2005). Möller (2005) presented a hypothetical scenario to 7th to 9th graders and college students, in which a third-party protagonist got equal performance on a math and a verbal test (e.g., got the same number of items correct on both tests) but only received praise (e.g., “Well done”) from the teacher on math performance. The participants were asked why the teacher had given different feedback for the same results and to estimate the teacher’s expectancy and the teacher’s beliefs about the protagonist’s effort and ability. Results showed paradoxical effects of praise, such that participants in both age groups attributed teachers’ praise to the students’ greater effort but lower ability. This study supported the third possibility mentioned above. However, it is unclear whether individuals would respond to receiving disproportionate praise themselves the same way as they did toward the hypothetical, third-party protagonist.

Additionally, this previous study only examined the situation when praise was given for math performance. Little is known about whether individuals would respond to praise for verbal performance the same way as for math performance. It is possible that participants may draw different conclusions when praise is given for verbal performance, because individuals believe that math performance relies more on natural talent than verbal performance (e.g., Gunderson et al., 2017). Specifically, because students’ fixed mindsets lead to the belief that high effort reflects low ability (Muenks & Miele, 2017), students’ fixed mindsets in math may lead students to interpret praise for math as reflecting their high effort and therefore low ability (a paradoxical effect). In contrast, students' belief that the verbal domain relies on effort may lead them to interpret praise as reflecting their high effort *and* high ability in the verbal domain, leading to positive effects of praise on verbal motivation. Investigating the impacts of praise in both math and verbal domains within a single study, as we do here, is crucial for detecting potential differences in students' responses to praise across domains.

Effects of Praise

In order to examine whether disproportionate praise can elicit dimensional comparisons across domains, it is important to understand how praise affects students' motivational beliefs. A large amount of work has suggested that adults can promote students' intrinsic motivation through certain forms of praise they give (Anderson et al., 1976; Mueller & Dweck, 1998; Zentall & Morris, 2010). For example, adults often give children *general* praise (e.g., "You did a good job" or "Awesome"), which is called *general* because it provides no explicit attributions for the outcome. Morris and Zentall (2014) found that general praise promoted task persistence and self-evaluations after failure in 5- to 6-year-old children.

In addition to general praise, teachers also provide more specific types of praise, including praise for students' effort or strategies (*process* praise; e.g., "You did a good job drawing the lines") and for students' personal traits or abilities (*person* praise; e.g., "You are very smart"). A previous study found that process praise enhanced college students' intrinsic motivation and perceived competence more than person praise (Haimovitz & Corpus, 2011). Although most prior work has focused on process and person praise, these two types of praise represent only a subset of the praise students receive, whereas the majority of praise that students receive is general praise (Gunderson et al., 2013; Pomerantz & Kempner, 2013). It is not clear what impacts general praise has on adolescents or adults. Therefore, more research on the effects of general praise, which is one of the goals in the current study, will provide insights into its influence on individuals' development of motivation.

It is important to note that general praise has slightly different definitions in different studies. Specifically, telling children "good job" was considered as process praise in some studies (e.g., Gunderson et al., 2013), but not others (e.g., Pomerantz & Kempner, 2013). In the current study, "good job" is considered to be general praise because it does not have explicit attributions for the outcome and is more neutral than process praise. For the purpose of simplicity, we referred to general praise as "praise" hereafter.

The Current Study

To summarize, building upon prior research on dimensional comparisons and the role of praise in shaping students' motivational beliefs, the current study investigated whether receiving disproportionate praise in different academic domains (math and verbal) induced students' dimensional comparisons. To test this, we asked students to complete math and verbal tasks that were selected to be challenging for their age group. We designed these tasks so that students' performance would be ambiguous, so that the reported performance scores and experimentally manipulated praise would be credible. To equate students' beliefs about their objective performance, all students were told that they received the same scores on the math and verbal tests. We then experimentally manipulated whether students received praise from the experimenter in the math domain, the verbal domain, or neither domain.

We hypothesized that receiving praise in one academic domain but not the other would result in higher motivational beliefs in the praised domain but lower motivational beliefs in the non-praised domain. Importantly, we assessed students' motivational beliefs (competence self-concepts and intrinsic values) in each domain before and after receiving praise to control for prior motivational beliefs when assessing the impact of praise. We chose to examine competence self-concepts and intrinsic values because these show stronger dimensional comparison effects than other motivational beliefs (Gaspard et al., 2018). Due to their high correlation, we expected competence self-concepts and intrinsic values to show similar patterns of effects.

We conducted the experiment in two age groups – 7^{th}- to 9^{th}-grade students and 1^{st}- to 5^{th}-year college students – to examine whether receiving disproportionate praise in different academic domains induces students' dimensional comparisons across these different developmental periods. We included 7^{th}- to 9^{th}-grade students because, according to prior literature, dimensional comparison effects begin to manifest in middle school, which is a time when students begin to make choices about which fields they want to pursue (e.g., Guo et al., 2017; Marsh et al., 2015; Putnick et al., 2019; Wan et al., 2021; Wolff et al., 2018; Zimmermann et al., 2018). However, to date, there is correlational but not experimental evidence to support the presence of dimensional comparison processes in 7^{th}- to 9^{th}-grade students. Moreover, most studies of dimensional comparison processes have focused on German students, wherein the educational system in Germany requires students to make decisions on which path (e.g., academic track or non-academic track) they are going to take in middle school, which may facilitate dimensional comparisons at an early age. In the United States, students generally do not need to make these decisions until late middle and high school. Indeed, a recent meta-analysis found that German students displayed more negative relations between math and verbal motivational beliefs compared to U.S. students (Wan et al., 2021). It is therefore important to examine dimensional comparison effects in 7^{th}- to 9^{th}-grade students in an experimental paradigm and in a different educational system.

We also examined college students, a population that has been shown to use dimensional comparisons in order to make career-relevant choices, such as choosing courses and college majors and applying to jobs (Nagy et al., 2006, 2008). Examining college students was also consistent with prior experimental studies on dimensional comparisons, which afforded the chance to directly compare results across studies. We did not expect to see differences between 7th- to 9th-graders and college students based on the consistent evidence of dimensional comparison effects in both age groups (e.g., Marsh & Yeung, 1998; Möller et al., 2009; Möller, 2005). However, due to the limited experimental research on how students in different ages interpret praise across domains, we did not exclude the possibility of seeing different patterns of results across age groups. If disproportionate praise does stimulate dimensional comparisons in 7th- to 9th-graders and college students, this may inform parents and teachers about how to provide praise that supports students' motivation across domains during different points in development.

Method

Participants

We recruited 7th- to 9th-graders and college students. We conducted a power analysis using the software program G*Power, with the goal to obtain .80 power to detect a small to medium effect size f = 0.20 (estimated from Pohlmann & Möller, 2009, study 3) at the standard .05 alpha error probability. A total sample size of 108 participants would be enough to detect the interaction effects between conditions and domains. Therefore, we aimed to test 108 participants per age group who completed both sessions.

Seventh- to Ninth-Grade Students

Participants were 7th- to 9th-grade students (*N* = 108, 46 girls; 39 seventh graders, 41 eighth graders, 28 ninth graders; age: *M* = 13.57 years, *SD* = 0.88). We recruited students by posting advertisements on social media and research communities (e.g., https://childrenhelpingscience.com). We asked parents whether they and their child were living in the U.S. at the time of testing. Ninety-two participants were living in the U.S., 14 were not living in the U.S., and two did not provide such information. Most parents reported their children's race and ethnicity (*n* = 105): 84.8% White, 2.9% Asian or Asian American, 2.9% Black or African American, and 9.5% multiple races. Parents' education ranged from less than high school to a graduate degree, with an average of 17.38 years of education (*SD* = 1.16 years, where 16 years is a 4-year college degree, *n* = 104). Family income ranged from less than $15,000 per year to more than $100,000 per year, with a mean of $88,250 (*SD* = $21,330, *n* = 87).

An additional 23 7th to 9th-grade students began the study but were excluded from our analytic sample for not having usable data for both sessions[1]. These included 21 children who did not complete the second session, 1 child whose second session data was excluded because of experimenter error (he was given 10 minutes to complete the reasoning tests), and 1 child whose second session data was excluded for having too many days between session 1 and session 2. After these exclusions, the final analytic sample for the 7th-9th graders was 108.

---

[1] The participants who only completed the first session of the study did not significantly differ on any of the baseline measures than participants who completed both sessions for the children and adult samples, except for self-reported math and verbal domain grades for children. The children who dropped out had significantly lower self-reported math grades (*M* = 3.24 , *SD* = .77) than children who completed both sessions (*M* = 3.69, *SD* = .46, $F(1,128) = 13.87$, $p < .001$, $\eta^2 = .098$) Children who only completed the first session also had significantly lower self-reported grades for the verbal domain (*M* = 3.37, *SD* = .73) than the children who completed both study sessions (*M* = 3.74, *SD* = .38, $F(1,129) = 12.02$, p <.001, $\eta^2 = .085$).

College Students

Participants were college undergraduates (*N* = 109, 89 women). We recruited college students through a northeastern U.S. university's psychology research participation system and by posting advertisements on social media. When participants signed up for our study, we asked them to use a school-issued email address to further confirm that they were in college. Eligible participants were students who were born between 1995 and 2002 (ages 18 to 25 years), had no difficulty understanding English, and had a desktop or laptop computer with internet access. We also asked whether they were currently living in the U.S.; 105 were living in the U.S. and four were not living in the U.S. (although three of those four were students in U.S. colleges). All participants reported their year in college: 42.2% in the 1^{st}-year, 23.9% in the 2^{nd}-year, 15.6% in the 3^{rd}-year, 16.5% in the 4^{th}-year, and 0.9% in the 5^{th}-year. Most participants self-reported their race (*n* = 107): 58.9% White, 21.5% Asian or Asian American, 10.3% Black or African American, 5.6% multiple races, and 3.7% other races.

An additional 27 college students began the study but were excluded from these analyses for not having usable data from both sessions. These included 26 participants who did not complete the second session and 1 participant whose second session data was excluded because he noted that he was aware that the scores for the reasoning tests were fake. After these exclusions, the final analytic sample for the college students was 109.

Materials

Math Reasoning Test

Seventh- to ninth-grade students. The math reasoning test consisted of 15 multiple-choice questions, aiming to test participants' math reasoning skills (e.g., "Quantity A: $\sqrt{9+16}$, Quantity B: $\sqrt{9}+\sqrt{16}$." Multiple-choice responses: "A. Quantity A is greater; B. Quantity B is greater; C. The two quantities are equal; D. The relationship cannot be determined from the information given”). Those 15 questions were chosen from the practice test of the Independent School Entrance Exam, developed and administered by the Educational Records Bureau (ERB) (Educational Records Bureau, 2018). We selected this test to be challenging for seventh- to ninth-grade students, and to be similar in format to the college students’ math reasoning test that was used in prior research (described below).

College students. The math reasoning test consisted of 15 multiple-choice questions, aiming to test participants' math reasoning skills (e.g., “m, p, and x are positive integers and mp = x; Quantity A: m, Quantity B: x.” Multiple-choice responses: “A. Quantity A is greater; B. Quantity B is greater; C. The two quantities are equal; D. The relationship cannot be determined from the information given”). This test was used in a previous study examining dimensional and social comparisons in a U.S. college student sample, and demonstrated sufficient ambiguity to enhance the credibility of the feedback manipulation (Strickhouser & Zell, 2015).

Language Reasoning Test

Seventh- to ninth-grade students. The language reasoning test consisted of 15 multiple-choice questions, aiming to test participants' verbal reasoning skills (e.g., “Although many climbers have tried to scale Mount Everest, few have succeeded in reaching the ______of the mountain.” Multiple-choice responses: “A. core; B. perimeter; C. pinnacle; D. support”). Those 15 questions were chosen from the practice test of the Independent School Entrance Exam, developed and administered by the Educational Records Bureau (ERB) (Educational Records Bureau, 2018).

College students. The language reasoning test consisted of 15 multiple-choice questions, aiming to test participants' verbal reasoning skills (e.g., “Mathematicians have a distinctive sense of beauty: they strive to present their ideas and results in a clear and compelling fashion, dictated by __________ as well as logic." Multiple-choice responses: "A. caprice; B. aesthetics; C. obligation; D. methodologies; E, intellect"). To be consistent with the number of items in the math reasoning test, these 15 questions were selected from the original 20-item measure that was used in previous research with college students. These items have been shown to be ambiguous enough to enhance the credibility of the feedback manipulation (Strickhouser & Zell, 2015).

Math and Verbal Competence Self-Concepts

We measured math and verbal competence self-concepts (i.e. beliefs about one's competence) using a 3-item scale for each domain, adapted from previous studies (Eccles et al., 1993; Wigfield et al., 1997). The wording was parallel across the two domains, and participants responded using a 1 to 6 Likert scale. The three items used were 1) "How good in [math/reading and writing] are you? [1: not good at all, 6: very good]", 2) "How well do you expect to do in [reading and writing/math] this year? [1: not good at all well, 6: very well", and 3) "How good would you be learning something new in [reading and writing /math]? [1: not good at all good, 6: very good]"Responses were averaged within each domain.

Math and Verbal Intrinsic Values

Intrinsic value (i.e., enjoyment) was measured using two items for each domain, adapted from previous studies (Eccles et al., 1993; Wigfield et al., 1997). Participants responded using a 1 to 6 Likert scale. The two items were 1) "In general, I find working on [reading and writing/math] assignments [1: very boring, 6: very interesting]" and 2) "How much do you like doing [math/reading and writing]? [1: not at all, 6: very much]." Responses were averaged within each domain.

Self-Reported Grades and College Major

To check for condition differences in prior achievement, we asked participants to report their math and verbal class grades by answering one question in each domain: "Please choose an option below that represents what kind of grades you receive in your [math/reading and writing] related classes." Options: “A. Mostly As; B. Mostly As and Bs; C. Mostly Bs; D. Mostly Bs and Cs; E. Mostly Cs; F. Mostly Cs and Ds; G. Mostly Ds; H. Mostly below Ds” (Gunderson et al., 2017). The response “Mostly As” was coded as 4, “Mostly As and Bs” was coded as 3.5, and each further option was 0.5 lower than the previous option.

We also asked college students to report their major (n = 106). We categorized majors as part of science, technology, engineering, or math (STEM) major or a non-STEM major based on the STEM Designated Degree Program List compiled by Department of Homeland Security (DHS) of the United States (Department of Homeland Security, 2021). We observed that 56% of college students in our sample were STEM majors (e.g., computer science, neuroscience, biology, chemistry, and psychology) and 44% were non-STEM majors (e.g., music therapy, marketing, finance, sociology, pharmacy, and criminal justice).

Manipulation Check

To assess the effectiveness of the experimental manipulation, participants were asked three questions. The first two questions assessed their satisfaction with their test results in each domain: “How satisfied are you with your result in the [math/language] reasoning test?” The choices ranged from 1 (not at all satisfied) to 6 (satisfied). The third question asked how they perceived their respective performance: “How do you view your result in the math reasoning test in relation to your result in the language reasoning test?” The choices ranged from 1 (math is much better) to 5 (verbal is much better), with 3 as “Same” (Möller & Köller, 2001).

Procedure

First Session

After parents of 7th- to 9th-graders and college students consented and completed an optional demographic survey through REDCap, participants were randomly assigned to one of three conditions (praise for math, praise for verbal, or no praise control) and one of two orders (math items first or verbal items first). Then they received a link to the first session online survey, which they could complete at their own convenience. Participants first completed self-reported math and reading and writing class grades, and then a theories of intelligence measure[2]. Following that, they completed two example items from the math and verbal reasoning tests. We selected these two example items from the 15 reasoning questions of each domain, and included them in order to provide a point of reference for participants to subsequently report their math and verbal competence self-concepts and intrinsic values. Upon completion of session 1, REDCap automatically recorded the date and participants were prompted to schedule a Zoom appointment for the second session. Each participant received a $5 gift card as compensation for completing the first session.

Second Session

The second session was arranged 3 to 20 days after participants completed the first session (7th to 9th graders: $M = 9.38$ days; college students: $M = 8.60$ days). Trained experimenters conducted one-on-one Zoom meetings with participants. The experimenter first introduced themself as a scientist at a university, then explained that the aim of the study was to

---

[2] As a secondary exploratory aim, we assessed students' theories of intelligence to explore the possibility that the higher incremental beliefs students had, the weaker the dimensional comparison effects would be induced by disproportionate praise. However, we did not find any significant interactions of Condition x Theories of Intelligence on students' motivational beliefs in either age group. See Appendix B for the detailed instruments and relevant analyses and results.

test the quality and efficiency of two tests being developed for a [seventh- to 9th-grade/college] student sample. Then participants were asked to complete the math and verbal reasoning tests, each within 6 minutes. To increase the credibility of the subsequent feedback, the items were presented sequentially until the end of the 6 minutes, and we did not inform participants of the total possible number of reasoning questions.

After each test, the computer presented a score on the screen (i.e., "The system has converted your result into a 100-point scale. You scored in the 75 to 85 range"). Regardless of students' actual performance[3], students were shown the same score for both tests and in all conditions. If participants were in one of the two praise conditions, the experimenter gave praise after either the math score was presented, or after the verbal score was presented, and did not praise the other domain. The praise was formulated as follows: "Awesome! This is a really, really good score! You did a great job on this task." If participants were in the no praise control condition, they did not receive praise in either domain. At the end of both reasoning tests, participants answered questions regarding their math and verbal competence self-concepts and intrinsic values. Finally, participants answered the three manipulation check questions. After completing the items, participants were debriefed and received a $15 gift card as compensation for completing the second session.

---

[3] We also checked students' performance on the two reasoning tests. Seventh to ninth graders completed fewer math questions (M = 9.26, SD = 3.10) than verbal questions (M = 13.69, SD = 2.13, $p < .001$). Correspondingly, the percent accuracy out of the total 15 items was higher in the verbal domain (M = .62, SD = .22) than the math domain (M = .40, SD = .20, $p < .001$). However, the percent accuracy out of completed items was not significantly different between domains (math: M = .64, SD = .22; verbal: M = .67, SD = .20, p = .123). College students' performance followed the same pattern. College students answered fewer math questions (M = 9.43, SD = 2.77) than verbal questions (M = 11.42, SD = 2.74, $p < .001$). The percent accuracy out of 15 items was higher in the verbal domain (M = .32, SD = .19) than the math domain (M = .26, SD = .14, p = .001). However, the percent accuracy out of completed items was not significantly different between domains (math: M = .43, SD = .23; verbal: M = .41, SD = .20, p = .409).

The study procedures were approved under [ANONYMIZED] IRB protocol [ANONYMIZED], “The Development of Motivational Beliefs in Dimensional Comparisons.”

Transparency and Openness

We report how we determined our sample size, all data inclusion and exclusion criteria, all manipulations, and all measures in the study. All data, analysis code, and research materials are available at https://osf.io/bpq49/?view_only=0dbdfbb1f7684840a384b4252e8dd3ec. Data were analyzed using R 3.6.3 (R Core Team, 2020), and the packages tidyverse (Wickham et al., 2019), psych (Revelle, 2021), and ggplot2 (Wickham, 2009). This study’s design and its analysis were not pre-registered.

Results

Preliminary Analyses

Table 1 presents the means, standard deviations, and correlations among key variables for 7^{th}- to 9^{th}-graders and college students. We checked whether students had similar levels of academic performance across different conditions by conducting one-way ANOVAs on self-reported grades, separately for math and verbal domains. For both children and college students, self-reported grades in both domains were not significantly different across conditions (*p*s $\geq .101$). We also checked college students’ percentage of STEM majors across conditions by conducting a Pearson’s chi-square test. The number of students who were in a STEM major did not significantly differ by condition ($\chi^2(2,106) = 1.84$, $p = .398$).

Table 1

Descriptive Statistics and Correlations Among Key Variables

Variables	*M* (*SD*)	1	2	3	4	5	6	7
Seventh- to ninth-grade students (n = 108)								
First session (before praise manipulation)								
1. Math competence self-concepts	4.97 (0.76)							
2. Verbal competence self-concepts	5.11 (0.68)	.17						
3. Math intrinsic values	4.01 (1.41)	.64***	.00					
4. Verbal intrinsic values	3.99 (1.31)	-.07	.57***	.09				
Second session (after praise manipulation)								
5. Math competence self-concepts	4.81 (0.82)	.76***	.12	.52***	-.11			
6. Verbal competence self-concepts	4.88 (0.77)	.07	.63***	-.15	.49***	.20*		
7. Math intrinsic values	4.10 (1.32)	.60***	.00	.88***	.05	.57***	-.14	
8. Verbal intrinsic values	4.21 (1.18)	-.06	.50***	.02	.81***	-.16	.52***	.01
College students (n = 109)								
First session (before praise manipulation)								
1. Math competence self-concepts	4.59 (0.91)							
2. Verbal competence self-concepts	4.87 (0.83)	.07						
3. Math intrinsic values	3.88 (1.39)	.62***	-.23*					
4. Verbal intrinsic values	3.73 (1.40)	-.18	.63***	-.23*				
Second session (after praise manipulation)								
5. Math competence self-concepts	4.57 (0.75)	.79***	-.02	.59***	-.19*			
6. Verbal competence self-concepts	4.72 (0.65)	-.01	.74***	-.23*	.60***	.04		
7. Math intrinsic values	4.11 (1.17)	.56***	-.23*	.88***	-.22*	.64***	-.20*	
8. Verbal intrinsic values	3.89 (1.21)	-.18	.66***	-.20*	.89***	-.19*	.67***	-.17

Note. Competence self-concept and intrinsic value scores ranged from 1 to 6.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Manipulation Checks

We checked our praise manipulation by conducting a two-way mixed effects ANOVA on satisfaction scores, with condition (praise for math, praise for verbal, and no praise control) as a between-subjects factor and domain (math, verbal) as a within-subjects factor. For both children and college students, we found a significant main effect of condition (children: $F(2,105) = 6.39$, $p = .002$, partial $\eta^2 = .109$; college students: $F(2,106) = 3.38$, $p = .038$, partial $\eta^2 = .060$). Neither children nor college students showed a main effect of domain (children: $F(1,105) = .79$, $p = .378$, partial $\eta^2 =.007$; college students: $F(1,106) = .34$, $p = .563$, partial $\eta^2 =.003$) nor a Domain x Condition interaction (children: $F(2,105) = 1.21$, $p = .304$, partial $\eta^2 =.022$; college students: $F(1,106) = 2.76$, $p = .068$, partial $\eta^2 =.049$). We further conducted planned pairwise comparison between conditions using Holm-Bonferroni corrections. For children, satisfaction scores were lower in the no praise control condition (estimated marginal $M = 3.89$, $SD = 1.06$) than both the praise for math condition (estimated marginal $M = 4.53$, $SD = 0.90$, adjusted $p < .001$), and the praise for verbal condition (estimated marginal $M = 4.53$, $SD = 0.99$, adjusted $p < .001$). Similarly, for college students, satisfaction scores were lower in the no praise control condition (estimated marginal $M = 3.85$, $SD = 1.08$) than both the praise for math condition (estimated marginal $M = 4.30$, $SD = 0.79$, adjusted $p = .013$), and the praise for verbal condition (estimated marginal $M = 4.35$, $SD = 1.13$, adjusted $p = .010$). Though we did not find evidence for higher satisfaction specifically in the praised domain, the overall higher satisfaction scores across both domains in the two praise conditions compared to the control condition still suggested that participants perceived praise as positive feedback which affected their overall experience.

Our other manipulation checks asked students to compare their perceived performance across the two domains, where higher scores indicated that they perceived their verbal performance as better. For children, the perceived performance score did not significantly differ between conditions (praise for math: $M = 3.03$, $SD = 0.57$; praise for verbal: $M = 2.97$, $SD = 0.72$; no praise control: $M = 3.06$, $SD = 0.81$), and were close to the scale midpoint of "3", indicating that children tended to interpret their performance as equal for both domains. For college students, however, the perceived performance score was significantly lower in the praise for math condition ($M = 2.73$, $SD = 0.56$) than the praise for verbal condition ($M = 3.21$, $SD = 0.65$, $p = .002$), and the control condition ($M = 3.18$, $SD = 0.76$, $p = .004$). These results showed that college students who received praise in math tended to think they had better math performance than verbal, compared to college students who received praise in the verbal domains and college students who did not receive praise. This suggested that our praise manipulation was effective, at least in the praise for math condition, in affecting college students' perception about their test scores.

Main Analyses

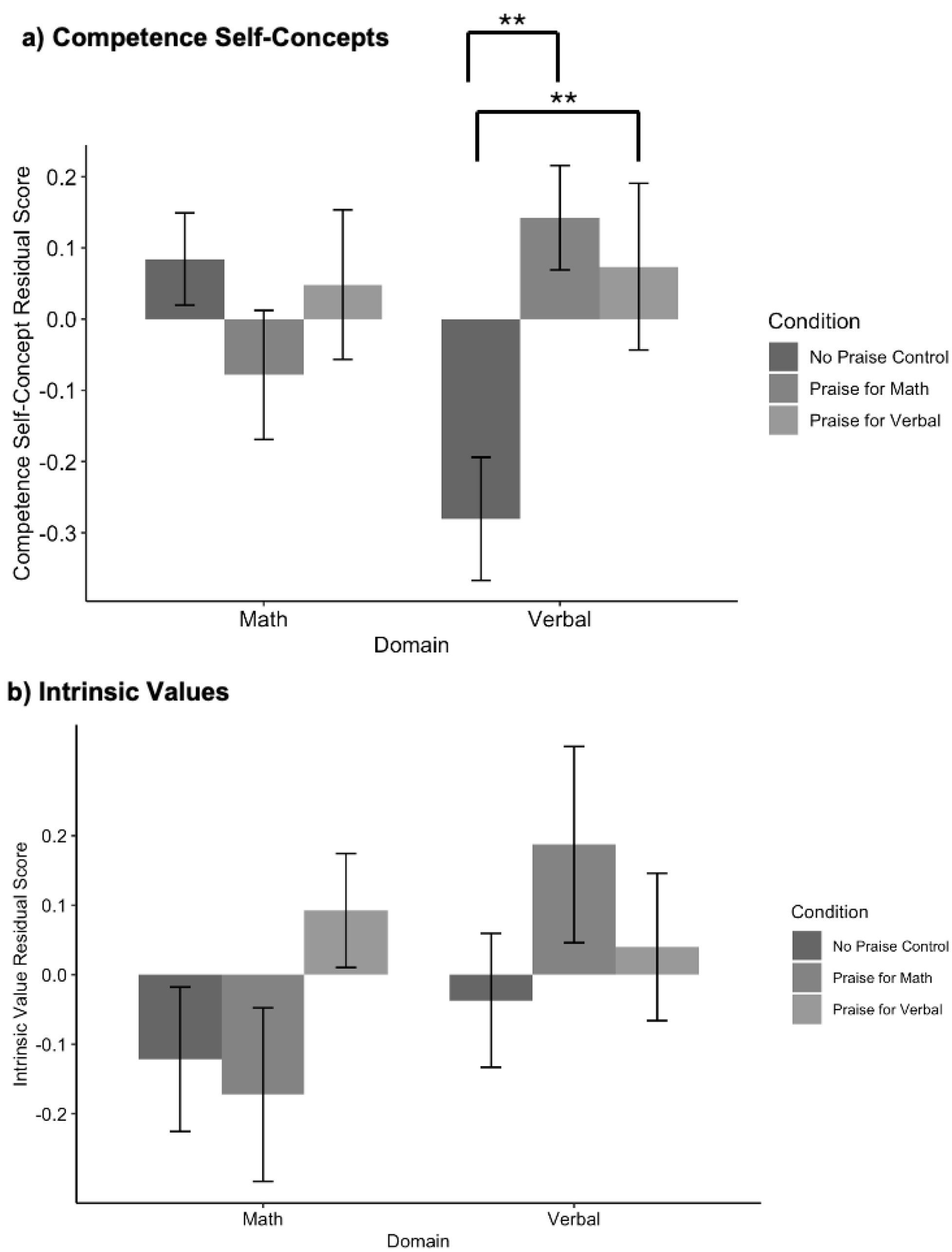
We hypothesized that receiving praise in one domain would result in higher motivational beliefs in that domain but lower motivational beliefs in the contrasting domain, compared to the control condition. To account for participants' motivational beliefs prior to the experimental manipulation, we computed competence self-concept and intrinsic value scores adjusted for their corresponding scores in the first session (consistent with the analytical method used in Pohlmann & Möller, 2009), across all domains and conditions, separately in each age group. These unstandardized residual scores were used as the dependent variables in the following analyses. See Appendix A for raw scores of competence self-concepts and intrinsic values, by condition and session.

Seventh- to Ninth-Grade Students

To test our main hypotheses within 7th to 9th-graders, we first conducted 3 × 2 mixed-effects ANOVAs using condition (praise for math, praise for verbal, and no praise control) as the between-subject variable, and domain (math and verbal) as the within-subject variable. We conducted two separate models with competence self-concept and intrinsic values as the DVs; see Figures 1a and 1b. For competence self-concepts, we found a significant Condition × Domain interaction ($F(2,105) = 7.73$, $p < .001$, partial $\eta^2 = .128$). However, we did not find a significant main effect of Condition ($F(2,105) = 1.24$, $p = .295$, partial $\eta^2 = .023$) or Domain ($F(1, 105) = .42$, $p = .518$, partial $\eta^2 = .004$, see Figure 1a).To interpret this interaction and directly test our hypothesis that receiving praise in a domain would result in higher competence self-concepts in that domain than not receiving praise, we conducted pairwise comparisons between conditions within each domain (3 comparisons within each domain) with Holm-Bonferroni corrections. Students' math competence self-concept residual scores did not significantly differ across the three conditions, adjusted *p*s > .05. Regarding verbal competence self-concepts, students had significantly lower verbal competence self-concept residual scores in the no praise condition ($M = -.28$, $SD = .51$) than both the praise for math condition ($M = .14$, $SD = .43$, adjusted $p = .003$) and the praise for verbal condition ($M = .07$, $SD = .72$, adjusted $p = .010$). The difference between the praise for math and praise for verbal conditions was not significant (adjusted $p = .611$). These results provided some evidence for our hypothesis, such that praising for verbal performance yielded higher verbal competence self-concepts than the control condition. However, we did not see that praise in one domain decreased competence self-concepts in the contrasting domain. Instead, contrary to our hypothesis, praising students' math performance resulted in higher verbal competence self-concepts than no praise.

Figure 1

Seventh- to Ninth-Grade Students' Mean Motivational Belief Residual Scores by Domain and Condition



Note. Error bars indicate standard errors.
* $p < .05$. ** $p < .01$. *** $p < .001$.

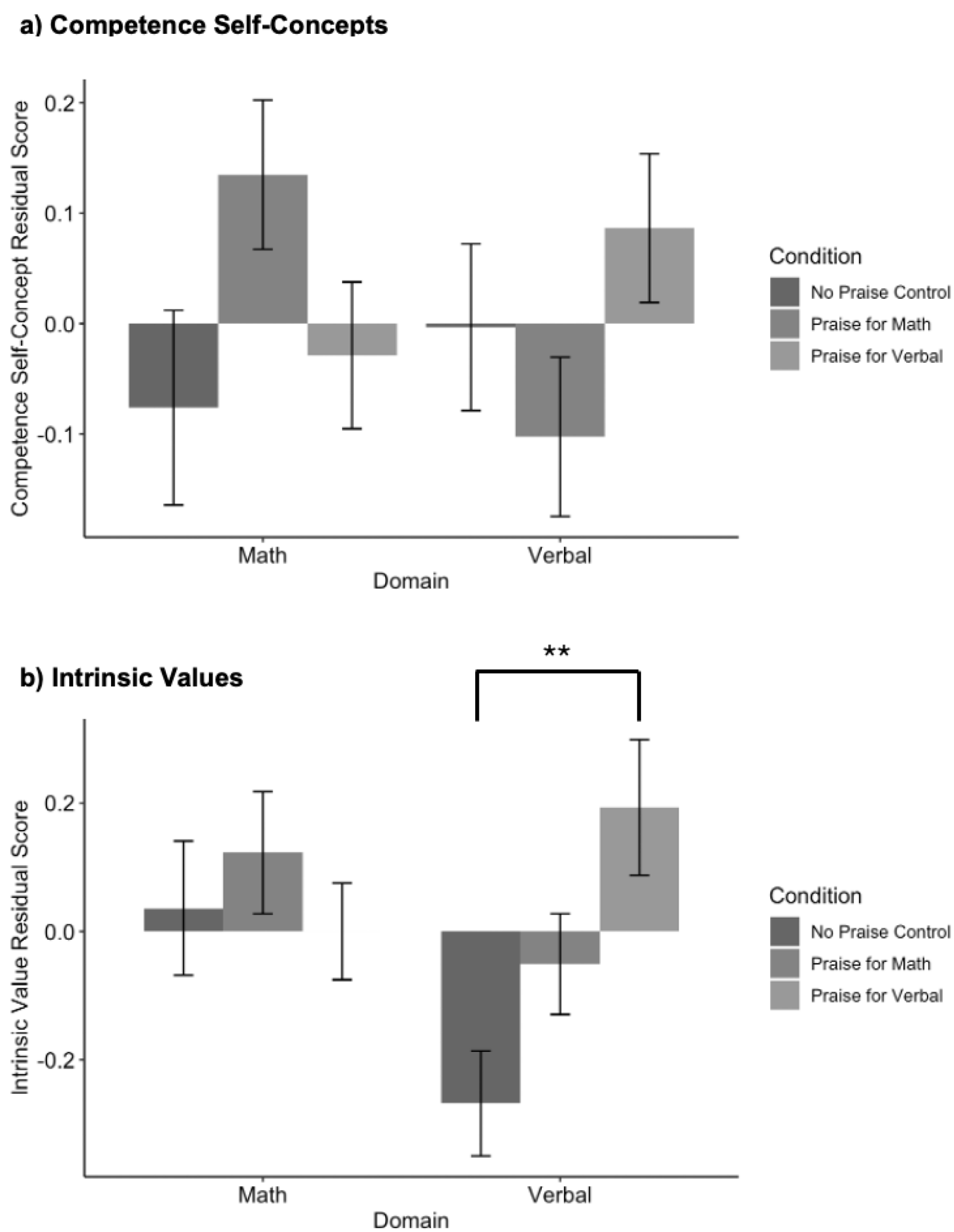
Finally, for intrinsic values, we did not find any significant effects in the 3 × 2 ANOVA (see Figure 1b). There was no significant main effect of Condition ($F(2,105) = .82$, $p = .446$, partial $\eta^2 = .015$), Domain ($F(1,105) = 2.33$, $p = .130$, partial $\eta^2 = .022$), or a Condition x Domain interaction ($F(2,105) = 2.02$, $p = .138$, partial $\eta^2 = .037$). We further conducted pairwise comparisons between conditions within each domain (3 comparisons within each domain) to directly test our hypothesis. Results showed no condition differences within either the math domain (adjusted *p*s ≥ .219) or the verbal domain (adjusted *p*s ≥ .179). Thus, we did not find evidence that praise impacted students' intrinsic values in either domain.

College Students

To test our main hypotheses within college students, we conducted 3 × 2 mixed-effects ANOVAs using condition and domain as factors, separately for competence self-concepts and intrinsic values as the dependent variables (two models). For competence self-concepts, we found a significant Condition × Domain interaction ($F(2,106) = 4.74$, $p = .011$, partial $\eta^2 = .082$, see Figure 2a). However, we did not find significant main effects of Condition ($F(2,106) = .38$, $p = .684$, partial $\eta^2 = .007$) or Domain ($F(1,106) = .10$, $p = .748$, partial $\eta^2 = .001$). To follow up on the significant Condition x Domain interaction, we conducted pairwise comparisons between conditions within each domain (3 comparisons within each domain). Results showed that math competence self-concept residual scores did not differ across conditions (adjusted *p*s ≥ .142). Regarding verbal competence self-concept residual scores, we also did not find condition differences (adjusted *p*s ≥ .073). Although the pairwise comparisons across conditions were not significant, the presence of a significant Condition x Domain interaction and the directional effects across conditions (see Figure 2b) were consistent with our hypothesis that praise in one domain would increase self-concepts in that domain and decrease self-concepts in the contrasting domain.

Figure 2

College Students' Mean Motivational Belief Residual Scores by Domains and Conditions



Note. Error bars stand for standard errors.
* p <.05. ** p < .01. *** p < .001.

Regarding intrinsic values, a 3 × 2 ANOVA revealed a significant Condition × Domain interaction ($F(2,106) = 4.40$, $p = .015$, partial $\eta^2 = .077$, see Figure 2b). However, we did not find significant main effects of Condition ($F(2,106) = 2.58$, $p = .081$, partial $\eta^2 = .046$) or Domain ($F(1,106) = 1.85$, $p = .176$, partial $\eta^2 = .017$). Pairwise comparisons between conditions within the math domain showed that intrinsic value residual scores did not differ across conditions (adjusted *p*s = 1). Within the verbal domain, the intrinsic value residual scores in the praise for verbal condition ($M = .19$, $SD = .61$) were significantly higher than those in the no praise control condition ($M = -.27$, $SD = .51$, adjusted $p = .001$). Residual verbal intrinsic value scores were intermediate after praise for math ($M = -.05$, $SD = .48$), and did not significantly differ from the other conditions (adjusted $ps \geq .116$). This indicates that praise for verbal performance increased students' verbal intrinsic values, partially consistent with our hypotheses.

Discussion

Here, we tested whether praising students only for their math performance, or only for their verbal performance, would lead to higher motivational beliefs in the praised domain, but lower motivational beliefs in the non-praised domain. In 7^{th} to 9^{th} graders, consistent with our hypothesis, praising students' verbal performance led to higher verbal competence self-concepts than receiving no praise (the control condition). We were surprised to find that praising 7^{th} to 9^{th} graders' math performance led to *higher* verbal competence self-concepts than the control condition, a dimensional comparison effect in the opposite direction than we had predicted. Also unexpectedly, praising 7^{th} to 9^{th} graders' math or verbal performance did not impact their math competence self-concepts relative to the control condition. Finally, our praise manipulations did not impact 7^{th} to 9^{th} graders' intrinsic values.

Among college students, we found that praise impacted students' competence self-concepts differently across the math and verbal domains, indicated by a significant Condition × Domain interaction. Although the follow-up pairwise comparisons were not significant, descriptively, receiving praise led to college students having higher competence self-concepts in the praised domain than after no praise. Consistent with this descriptive pattern, in our manipulation check, college students reported perceiving their math performance as significantly better than their verbal performance when receiving praise on math performance. In terms of intrinsic values, we found that college students had higher verbal intrinsic values after receiving praise for verbal performance, compared to the control condition, supporting our hypothesis.

Taken together, we believe that our praise manipulations were, though simple and brief, still impactful in affecting students' domain-specific motivational beliefs. It is intriguing that 7th to 9th graders responded differently to receiving praise for their math performance than for their verbal performance, in terms of their within-domain and across-domain competence self-concepts. Specifically, receiving praise for math performance led to no significant change in 7th to 9th-graders' math competence self-concepts but led to higher verbal competence self-concepts. In contrast, when receiving praise on verbal performance, 7th to 9th-graders' verbal competence self-concepts were significantly higher than the control condition, as we hypothesized. This suggests that they perceived the praise for their verbal performance as an indicator of the experimenters' belief in their high ability, whereas they may have perceived the praise for their math performance as an indicator of the experimenters' belief in their high effort. In the verbal domain, students may view the link between effort and ability as positive, such that hard work improves verbal abilities. This possibility is partially supported by a prior study that showed that students believed reading and writing involves more hard work and effort than math, which they viewed as involving more fixed and unchangeable ability (Gunderson et al., 2017).

The null effect of praise for math on 7th to 9th-graders' math competence self-concepts might be explained by some students' belief that ability and effort in math are inversely related (for a review, see Amemiya & Wang, 2018). Students may view praise for math as reflecting their high effort and therefore lower ability, consistent with a prior study by Möller (2005). However, in our study, students' math competence self-concepts were directionally, but not significantly, lower after praise for math than in the control condition; this non-significant result may indicate that students may respond heterogeneously to praise for math at this age. Future research might benefit from directly measuring students' beliefs about the ability-effort relation in different domains. Such studies would help to gain insight into why students respond to feedback differently across domains and how teachers and parents could better motivate students.

Unlike 7th to 9th graders, college students showed descriptively higher motivational beliefs in the praised domains and lower beliefs in the non-praised domains, and a significant Condition × Domain interaction, consistent with our predictions. Prior research suggests that dimensional comparison effects are stronger at older ages than younger ages (Möller et al,. 2020). This may help to explain why college students showed the expected dimensional comparison effects on self-concepts whereas 7th to 9th graders did not. In other words, even though both age groups are expected to engage in dimensional comparison processes (e.g., Marsh & Yeung, 1998; Möller et al., 2009; Möller & Köller, 2001), stronger dimensional comparisons may be required to show effects in a short-term experimental study.

Additionally, college students' intrinsic values seemed to be more malleable as a result of praise than 7th to 9th graders', at least in the verbal domain, in which college students' intrinsic values increased after praise. It could be that college students are focused more on whether they like doing something as they consider their career path, whereas 7th to 9th graders are more likely to value their abilities more than their individual interests. This possibility is aligned with the performance-oriented school culture in the middle school years (Midgley et al., 1995). Altogether, these distinct patterns highlight the potentially different impacts of praise towards students of various ages. Including these two developmental age groups was a strength of the current study and these preliminary developmental findings should be replicated in future studies, such as cross-sectional studies with more age groups and longitudinal studies that can better examine how the impacts of praise may change across development.

Limitations and Future Directions

The current studies have several limitations that should be considered. First, though we chose math and verbal reasoning test items from the same source, we did not quantitatively equate the difficulty levels across domains. Both 7th to 9th graders and college students answered fewer items in the math test than in the verbal test within the same amount of time (see Footnote 1). Though the percent accuracy out of completed items was not statistically different between the two domains, students may still have perceived the math test as more difficult. This may have contributed to their different interpretations of the experimenter's praise when given in the math domain compared to the verbal domain. Given our random-assignment design, we were able to robustly examine the effect of praise across conditions; however, these potentially different difficulty perceptions may have produced some bias in comparisons of the effect of praise between domains. For example, if students perceived the math test as being harder than the verbal test, when they received the same score range from the computer on both math and verbal performance, students may have inferred that they either had higher abilities or expended greater effort on the math test (Meyer & Hallermann, 1974). Individual differences in how students think about task difficulty may further affect how they respond to praise. Therefore, it would be helpful for future studies to control or experimentally manipulate the difficulty level of the test and investigate how students' perceptions of difficulty interact with their responses to praise.

Second, in order to manipulate praise experimentally, experimenters gave praise to students. However, this design may affect the generalizability of the current findings because the experimenters were individuals the students had never met and were unlikely to encounter again. It is possible that the effect of praise would be larger when given by parents or teachers, with whom students have an established relationship and whose opinions are more important to them. Students also received praise over Zoom instead of in person in the current study. The effect of praise might have been stronger if students received it in a more naturalistic setting, such as a classroom. Thus, future studies may investigate the effects of praise from teachers and parents in a naturalistic, in-person setting to shed more light on how praise can influence domain-specific motivational beliefs.

Third, our manipulation check indicated that the praise manipulation may not have been as strong as it could have been. For example, for both age groups, we found that students reported being more satisfied with their scores if they received praise in either domain compared to not receiving praise. One assumption of the design of our study was that domain-specific praise would have domain-specific effects. However, receiving domain-specific praise might generalize to other academic areas, leading to a more general effect of praise on students' satisfaction. Additionally, when explicitly asked whether their performance was better in the math or verbal domain, 7th-9th graders did not report that their performance was significantly different across conditions, which suggests that their explicit views of their performance seemed to be driven by the equal scores they received on each test regardless of whether they received domain-specific praise. However, college students reported that their math performance was greater than their verbal performance in the praise for math condition only. This suggests that the manipulation successfully increased college students' explicitly-reported perception of their perceived performance for the praise in math condition but not for the praise in verbal condition. This difference between the domains could have been due to several factors, such as the difference in difficulty level between the two tests or students' perceptions about their abilities or competencies in math compared to the verbal domain.

Further, when considering the effects of condition on the manipulation check compared to the effects on students' competence self-concepts and intrinsic motivation, the findings suggest that the praise manipulation may have had different implicit versus explicit effects. For example, the praise manipulation may have been too subtle for most students to recognize and report on when explicitly asked about their satisfaction and perceived performance in each domain. However, praise may nevertheless have impacted students at a more implicit level, so that when asked about their overall competence self-concepts and intrinsic motivation in each domain, we see some effects of the praise manipulation. Future research should use a stronger praise manipulation to better communicate to students that one domain is being praised and to get confirmation at the explicit level that students clearly understand which domain is being praised to see if the main findings of the current study replicate.

It is also possible that the discrepancy between the praise manipulation and outcome measures made it more difficult to find effects. For example, the praise participants received was task specific (e.g., “You did a great job on this task”) whereas the outcome measures were person-specific (e.g., “How good in math are you?"). There is evidence that feeling accomplished on a math task does not always mean that one feels like they are good at math in general (e.g., Marsh et al., 2019). So praising participants’ performance on one task may have relatively weak effects on self-concepts in a domain such as math or verbal ability. Nevertheless, we found some significant effects in this paradigm, consistent with our expectation that task-specific praise can influence self-concepts. Future research could explore whether these effects would be stronger if there were a match between the praise manipulation and outcome measures (e.g., both task-specific or both person-specific).

There were also some limitations regarding the samples recruited in the current study. For example, the current study included both children and young adults who were living in and outside of the U.S. This was a result of the online recruitment methods used (e.g., social media advertisements and Children Helping Science website). Although the number of children (n = 14) and young adults (n = 4) living outside of the U.S. in the current study was relatively small and there is evidence that dimensional comparison effects generalize across countries (e.g., Möller et al., 2009), future research should be conducted to see if the findings in the current study replicate with samples from different countries. The college student sample also had an overrepresentation of participants who identified as female (82%). Therefore, future research is needed to determine if certain participant characteristics, such as gender, moderate the findings of the current study.

Conclusion

We experimentally manipulated praise in the present study and found that 7^{th} to 9^{th} graders and college students showed distinct patterns of motivational beliefs after receiving praise for their verbal performance versus their math performance. When students received praise for verbal performance, this had positive effects on verbal motivational beliefs in both age groups. When students received praise for math performance, this induced dimensional comparisons in 7th to 9th-graders (but not in college students), as indicated by higher verbal motivational beliefs. However, students' math motivational beliefs were not affected by receiving praise in either domain. Overall, these results offered a promising first look at the role of disproportionate praise in students' domain-specific motivational beliefs. We hope to inspire future research to further explore the development of domain-specific motivational beliefs, given their significance in affecting students' future goals and career paths.

References

Amemiya, J., & Wang, M.-T. (2018). Why effort praise can backfire in adolescence. *Child Development Perspectives*, *12*(3), 199–203. https://doi.org/10.1111/cdep.12284

Anderson, R., Manoogian, S. T., & Reznick, J. S. (1976). The undermining and enhancing of intrinsic motivation in preschool children. *Journal of Personality and Social Psychology*, *34*(5), 915–922. https://doi.org/10.1037/0022-3514.34.5.915

Barker, G. P., & Graham, S. (1987). Developmental study of praise and blame as attributional cues. *Journal of Educational Psychology*, *79*(1), 62–66. https://doi.org/10.1037/0022-0663.79.1.62

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

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, *15*(1), 1–40. https://doi.org/10.1023/A:1021302408382

Cvencek, D., Brečić, R., Gaćeša, D., & Meltzoff, A. N. (in press). Development of Math Attitudes and Math Self-Concepts: Gender Differences, Implicit–Explicit Dissociations, and Relations to Math Achievement. *Child Development*, *n/a*(n/a). https://doi.org/10.1111/cdev.13523

Department of Homeland Security. (2021). *STEM Designated Degree Program List*. https://www.ice.gov/sites/default/files/documents/stem-list.pdf

Dweck, C. S. (2008). *Mindset: The new psychology of success*. Random House.

Dweck, C. S., Davidson, W., Nelson, S., & Enna, B. (1978). Sex differences in learned helplessness: II The contingencies of evaluative feedback in the classroom and III An experimental analysis. *Developmental Psychology*, *14*(3), 268–276. https://doi.org/10.1037/0012-1649.14.3.268

Eccles, J. S. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist*, *44*(2), 78–89. https://doi.org/10.1080/00461520902832368

Eccles, J. S. (2015). Gendered socialization of STEM interests in the family. *International Journal of Gender, Science and Technology*, *7*(2), 116–132.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J., & Midgely, C. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75–146). Freeman.

Eccles, J. S., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, *64*(3), 830–847. JSTOR. https://doi.org/10.2307/1131221

Educational Records Bureau. (2018). *ISEE Practice Test*. https://www.erblearn.org/parents/isee-preparation

Floress, M. T., Jenkins, L. N., Reinke, W. M., & McKown, L. (2018). General education teachers' natural rates of praise: A preliminary investigation. *Behavioral Disorders*, *43*(4), 411–422. https://doi.org/10.1177/0198742917709472

Gaspard, H., Wigfield, A., Jiang, Y., Nagengast, B., Trautwein, U., & Marsh, H. W. (2018). Dimensional comparisons: How academic track students' achievements are related to their expectancy and value beliefs across multiple domains. *Contemporary Educational Psychology*, *52*, 1–14. https://doi.org/10.1016/j.cedpsych.2017.10.003

Gladstone, J. R., Häfner, I., Turci, L., Kneißler, H., & Muenks, K. (2018). Associations between parents and students' motivational beliefs in mathematics and mathematical performance: The role of gender. *Contemporary Educational Psychology*, *54*, 221–234. https://doi.org/10.1016/j.cedpsych.2018.06.009

Gniewosz, B., Eccles, J. S., & Noack, P. (2014). Early adolescents' development of academic self-concept and intrinsic task value: The role of contextual feedback. *Journal of Research on Adolescence*, *25*(3), 459–473. https://doi.org/10.1111/jora.12140

Goetz, T., Frenzel, A. C., Hall, N. C., & Pekrun, R. (2008). Antecedents of academic emotions: Testing the internal/external frame of reference model for academic enjoyment. *Contemporary Educational Psychology*, *33*(1), 9–33. https://doi.org/10.1016/j.cedpsych.2006.12.002

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

Gunderson, E. A., Hamdan, N., Sorhagen, N. S., & D'Esterre, A. P. (2017). Who needs innate ability to succeed in math and literacy? Academic-domain-specific theories of intelligence about peers versus adults. *Developmental Psychology*, *53*(6), 1188–1205. https://doi.org/10.1037/dev0000282

Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles*, *66*(3–4), 153–166. https://doi.org/10.1007/s11199-011-9996-2

Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J. S., & Dicke, T. (2017). Extending expectancy-value theory predictions of achievement and aspirations in science: Dimensional comparison processes and expectancy-by-value interactions. *Learning and Instruction*, *49*, 81–91. https://doi.org/10.1016/j.learninstruc.2016.12.007

Haimovitz, K., & Corpus, J. H. (2011). Effects of person versus process praise on student motivation: Stability and change in emerging adulthood. *Educational Psychology*, *31*(5), 595–609. https://doi.org/10.1080/01443410.2011.585950

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hong, Y., Chiu, C., Dweck, C. S., Lin, D. M.-S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, *77*(3), 588–599. https://doi.org/10.1037/0022-3514.77.3.588

Hyland, F., & Hyland, K. (2001). Sugaring the pill: Praise and criticism in written feedback. *Journal of Second Language Writing*, *10*(3), 185–212. https://doi.org/10.1016/S1060-3743(01)00038-8

Lam, S., Yim, P., & Ng, Y. (2008). Is effort praise motivational? The role of beliefs in the effort–ability relationship. *Contemporary Educational Psychology*, *33*(4), 694–710. https://doi.org/10.1016/j.cedpsych.2008.01.005

Marsh, H. W., Abduljabbar, A. S., Parker, P. D., Morin, A. J. S., Abdelfattah, F., Nagengast, B., Möller, J., & Abu-Hilal, M. M. (2015). The internal/external frame of reference model of self-concept and achievement relations: Age-cohort and cross-cultural differences. *American Educational Research Journal*, *52*(1), 168–202. https://doi.org/10.3102/0002831214549453

Marsh, H. W., & Hau, K.-T. (2004). Explaining paradoxical relations between academic self-concepts and achievements: Cross-cultural generalizability of the internal/external frame of reference predictions across 26 countries. *Journal of Educational Psychology*, *96*(1), 56–67. https://doi.org/10.1037/0022-0663.96.1.56

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology, 111*(2), 331.

Marsh, H.W., & Yeung, A.S. (1998). Longitudinal Structural Equation Models of Academic Self-Concept and Achievement: Gender Differences in the Development of Math and English Constructs. *American Educational Research Journal, 35*(4), 705-738.

Meyer, W.-U., & Hallermann, B. (1974). The relationship between perceived ability and the intended expenditure of effort on easy and difficult tasks. *Archiv Für Psychologie*, *126*(1), 85–89.

Midgley, C., Anderman, E., & Hicks, L. (1995). Differences between elementary and middle school teachers and students: A goal theory approach. *The Journal of Early Adolescence*, *15*(1), 90–113. https://doi.org/10.1177/0272431695015001006

Möller, J. (2005). Paradoxical effects of praise and criticism: Social, dimensional and temporal comparisons. *British Journal of Educational Psychology*, *75*(2), 275–295. https://doi.org/10.1348/000709904X24744

Möller, J., & Köller, O. (2001). Dimensional comparisons: An experimental approach to the internal/external frame of reference model. *Journal of Educational Psychology*, *93*(4), 826–835. https://doi.org/10.1037/0022-0663.93.4.826

Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review*, *120*(3), 544–560. https://doi.org/10.1037/a0032459

Möller, J., Pohlmann, B., Köller, O., & Marsh, H.W. (2009). A Meta-Analytic Path Analysis of the Internal/External Frame of Reference Model of Academic Achievement and Academic Achievement and Academic Self-Concept. *Review of Educational Research, 79*(3), 1129-1167.

Möller, J., Zitzmann, S., Helm, F., Machts, N., & Wolff, F. (2020). A meta-analysis of relations between achievement and self-concept. *Review of Educational Research*, *90*(3), 376-419. https://doi.org/10.3102/0034654320919354

Morris, B. J., & Zentall, S. R. (2014). High fives motivate: The effects of gestural and ambiguous verbal praise on motivation. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00928

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, *75*(1), 33–52. https://doi.org/10.1037/0022-3514.75.1.33

Muenks, K., & Miele, D. B. (2017). Students' thinking about effort and ability: The role of developmental, contextual, and individual difference factors. *Review of Educational Research*, *87*(4), 707–735. https://doi.org/10.3102/0034654316689328

Musu-Gillette, L. E., Wigfield, A., Harring, J. R., & Eccles, J. S. (2015). Trajectories of change in students' self-concepts of ability and values in math and college major choice.

Educational Research and Evaluation, *21*(4), 343–370. https://doi.org/10.1080/13803611.2015.1057161

Nagy, G., Garrett, J., Trautwein, U., Cortina, K. S., Baumert, J., & Eccles, J. S. (2008). Gendered high school course selection as a precursor of gendered careers: The mediating role of self-concept and intrinsic value. In *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences* (pp. 115–143). American Psychological Association. https://doi.org/10.1037/11706-004

Nagy, G., Trautwein, U., Baumert, J., Köller, O., & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation*, *12*(4), 323–345. https://doi.org/10.1080/13803610600765687

Pesu, L., Aunola, K., Viljaranta, J., Hirvonen, R., & Kiuru, N. (2018). The role of mothers' beliefs in students' self-concept of ability development. *Learning and Individual Differences*, *65*, 230–240. https://doi.org/10.1016/j.lindif.2018.05.013

Pesu, L., Aunola, K., Viljaranta, J., & Nurmi, J.-E. (2016). The development of adolescents' self-concept of ability through grades 7-9 and the role of parental beliefs. *Frontline Learning Research*, *4*. https://jyx.jyu.fi/handle/123456789/51237

Pesu, L., Viljaranta, J., & Aunola, K. (2016). The role of parents' and teachers' beliefs in children's self-concept development. *Journal of Applied Developmental Psychology*, *44*, 63–71. https://doi.org/10.1016/j.appdev.2016.03.001

Pohlmann, B., & Möller, J. (2009). On the benefit of dimensional comparisons. *Journal of Educational Psychology*, *101*(1), 248–258. https://doi.org/10.1037/a0013151

Pomerantz, E. M., & Kempner, S. G. (2013). Mothers' daily person and process praise: Implications for children's theory of intelligence and motivation. *Developmental Psychology*, *49*(11), 2040–2046. https://doi.org/10.1037/a0031840

Putnick, D. L., Hahn, C.-S., Hendricks, C., Suwalsky, J. T. D., & Bornstein, M. H. (2019). Child, mother, father, and teacher beliefs about child academic competence: Predicting math and reading performance in European American adolescents. *Journal of Research on Adolescence*. https://doi.org/10.1111/jora.12477

R Core Team. (2020). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Revelle, W. (2021). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University. https://CRAN.R-project.org/package=psych

Seo, E., Shen, Y., & Alfaro, E. C. (2019). Adolescents' beliefs about math ability and their relations to STEM career attainment: Joint consideration of race/ethnicity and gender. *Journal of Youth and Adolescence*, *48*(2), 306–325. https://doi.org/10.1007/s10964-018-0911-9

Simpkins, S. D., Fredricks, J. A., & Eccles, J. S. (2012). Charting the Eccles' expectancy-value model from mothers' beliefs in childhood to youths' activities in adolescence. *Developmental Psychology*, *48*(4), 1019–1032. https://doi.org/10.1037/a0027468

Simpkins, S. D., Fredricks, J. A., & Eccles, J. S. (2015). The role of parents in the ontogeny of achievement-related motivation and behavioral choices: IV Child factors and parent belief models. *Monographs of the Society for Research in Child Development*, *80*(2), 65–84. https://doi.org/10.1111/mono.12160

Strickhouser, J. E., & Zell, E. (2015). Self-evaluative effects of dimensional and social comparison. *Journal of Experimental Social Psychology*, *59*, 60–66. https://doi.org/10.1016/j.jesp.2015.03.001

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, *92*(1), 144–151. https://doi.org/10.1037/0022-0663.92.1.144

Wan, S., Lauermann, F., Bailey, D. H., & Eccles, J. S. (2021). When do students begin to think that one has to be either a "math person" or a "language person"? A meta-analytic review. *Psychological Bulletin*, *147*(9), 867–889. https://doi.org/10.1037/bul0000340

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://doi.org/10.1007/978-0-387-98141-3

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J. A., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology*, *89*(3), 451–469. https://doi.org/10.1037/0022-0663.89.3.451

Wolff, F., Helm, F., Zimmermann, F., Nagy, G., & Möller, J. (2018). On the effects of social, temporal, and dimensional comparisons on academic self-concept. *Journal of Educational Psychology*, *110*(7), 1005–1025. https://doi.org/10.1037/edu0000248

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. https://doi.org/10.1016/j.jecp.2010.04.015

Zimmermann, F., Möller, J., & Köller, O. (2018). When students doubt their teachers’ diagnostic competence: Moderation in the internal/external frame of reference model. *Journal of Educational Psychology*, *110*(1), 46–57. https://doi.org/10.1037/edu0000196

Appendix A

Table A1

Means (SDs) of Motivational Beliefs, by Condition and Session

Variables	Condition		
	Praise for math	Praise for verbal	No praise control
Seventh- to ninth-grade students			
First session (before praise manipulation)			
Math competence self-concepts	4.84 (0.95)	5.24 (0.57)	4.80 (0.65)
Verbal competence self-concepts	5.10 (0.67)	5.08 (0.77)	5.16 (0.61)
Math intrinsic values	3.80 (1.39)	4.42 (1.36)	3.77 (1.44)
Verbal intrinsic values	3.94 (1.39)	4.00 (1.24)	4.03 (1.35)
Second session (after praise manipulation)			
Math satisfaction scores	4.51 (.89)	4.39 (1.05)	3.91 (1.12)
Verbal satisfaction scores	4.54 (.92)	4.66 (.91)	3.86 (1.00)
Perceived performance scores	3.03 (.57)	2.97 (.72)	3.06 (.81)
Math competence self-concepts	4.61 (1.00)	5.04 (0.74)	4.74 (0.65)
Verbal competence self-concepts	5.03 (0.63)	4.95 (0.88)	4.66 (0.74)
Math intrinsic values	3.83 (1.32)	4.58 (1.18)	3.86 (1.36)
Verbal intrinsic values	4.30 (1.15)	4.20 (1.19)	4.14 (1.24)
College students			
First session (before praise manipulation)			
Math competence self-concepts	4.61 (0.84)	4.45 (1.07)	4.68 (0.83)
Verbal competence self-concepts	4.85 (0.74)	4.94 (0.66)	4.83 (1.02)
Math intrinsic values	3.97 (1.17)	3.83 (1.32)	3.83 (1.65)
Verbal intrinsic values	3.54 (1.42)	3.88 (1.31)	3.79 (1.47)
Second session (after praise manipulation)			
Math satisfaction scores	4.41 (.80)	4.27 (1.07)	3.74 (1.14)
Verbal satisfaction scores	4.19 (.78)	4.42 (1.20)	3.95 (1.02)
Perceived performance scores	2.73 (.56)	3.21 (.65)	3.18 (.76)

Math competence self-concepts	4.70 (0.66)	4.44 (0.74)	4.54 (0.83)
Verbal competence self-concepts	4.61 (0.61)	4.86 (0.53)	4.70 (0.77)
Math intrinsic values	4.24 (0.94)	4.02 (1.15)	4.05 (1.38)
Verbal intrinsic values	3.74 (1.24)	4.24 (1.09)	3.72 (1.26)

Appendix B

Theories of Intelligence and Dimensional Comparisons

To better understand the relation between praise and students' motivational beliefs, we explored whether this relation was moderated by students' theories of intelligence. *Theories of intelligence* (TOIs) refer to individuals' beliefs about whether abilities are fixed or malleable and whether performance can be improved by effort or not (TOIs; Dweck, 2006). Individuals who hold an entity theory of intelligence (believing ability is fixed) tend to be more performance-oriented, less persistent, and more frustrated when facing difficulties, compared to individuals who hold an incremental theory of intelligence (viewing ability as malleable) (e.g., Blackwell et al., 2007; Hong et al., 1999). We expected that students who hold an entity theory would engage in more dimensional comparisons and show stronger negative effects on contrasting domains as they believe in the power of intelligence, pay more attention to ability, and are more performance-oriented. In contrast, students who hold an incremental theory may be less likely to engage in dimensional comparisons and less likely to show negative effects in dimensional comparisons because they believe that they can improve their performance by devoting more effort. Therefore, we expected that when students receive disproportionate praise, entity theorists would be more likely to engage in dimensional comparison processes and thus show stronger effects of praise on motivation than incremental theorists. Therefore, investigating the role of TOIs as a potential moderator of students' responses to disproportionate praise may enrich our understanding of individual differences within dimensional comparisons.

Theories of intelligence (TOIs) were measured using 3 items used in prior studies (Hong et al., 1999). The items were "You have a certain amount of intelligence and you really can't do much to change it"; "Your intelligence is something about you that you can't change very much"; and "You can learn new things, but you can't really change your basic intelligence." Participants were asked to respond on a 6-point Likert scale (1: strongly disagree to 6: strongly agree). All items were reverse-coded, so that the higher the participants' scores, the more they reported that intelligence was malleable (incremental beliefs). Thus, we refer to this measure as *incremental theory of intelligence* in our analyses to remind readers of this scoring rule. Items were averaged together to be used in data analyses.

Seventh- to ninth-graders' incremental theories of intelligence ($M = 4.52$, $SD = 1.26$), were significantly correlated with their verbal competence self-concepts measured in the first session ($r(106) = .20$, $p = .038$), but not with any other measures of self-concept or intrinsic value (ps > .05). Among college students, incremental theories of intelligence ($M = 4.31$, $SD = 1.13$) were not significantly correlated with self-concepts or intrinsic values in math or reading (ps > .05).

To examine whether individual differences in students' theories of intelligence moderated dimensional comparison processes after receiving disproportionate praise, we compared each experimental condition with the control condition. We conducted ANCOVAs with either competence self-concept or intrinsic value as the outcome, separately for the math and verbal domains. To control Type I error across the 8 models, we used the Bonferroni correction (.05/8) and considered $p < .006$ to be significant. In each model, we included condition as a factor and TOIs as a covariate. Specifically, we were looking for significant interactions between condition and TOIs, which would indicate that responses to praise differed as a function of students' TOIs. After correcting for multiple comparisons, we did not find any significant interactions of Condition x TOI on students' motivational beliefs in either age group.