Using Wise Interventions to Motivate Deliberate Practice

Lauren Eskreis-Winkler
University of Pennsylvania

Elizabeth P. Shulman
Brock University

Victoria Young
University of Pennsylvania

Eli Tsukayama
University of Southern California

Steven M. Brunwasser
Vanderbilt University

Angela L. Duckworth
University of Pennsylvania

Deliberate practice leads to world-class excellence across domains. In the current investigation, we examined whether psychologically "wise" interventions targeting expectancies and values—stock antecedents of ordinary effortful behaviors—could motivate nonexperts to engage in deliberate practice and improve their achievement. As a preliminary, we developed and validated a novel task measure of deliberate practice and confirmed its association with (a) expectancy-value beliefs, and (b) achievement in the nonexpert setting (Study 1). Next, across 4 longitudinal, randomized-controlled, field experiments, we intervened. Among lower-achievers, wise deliberate practice interventions improved math performance for 5th and 6th graders (Study 2), end-of-semester grades for undergraduates (Study 3), and end-of-quarter grades for 6th graders (Study 4); the same pattern of results emerged in end-of-quarter grades for 7th graders (Study 5). Following the intervention, expectancy-value beliefs and deliberate practice improved for 1 month (Study 4), but not 4 (Study 5). Treatment proved beneficial over and above 2 control conditions: 1 that taught standard study skills (Studies 2 and 3), and 1 that discussed deep interests, generalized motivation, and high achievement (Studies 4 and 5). Collectively, these findings provide preliminary support for the heretofore untested hypothesis that deliberate practice submits to the same laws that govern typical forms of effortful behavior, and that wise interventions that tap into these laws can spur short-term gains in adaptive beliefs, deliberate practice, and objectively measured achievement.

Keywords: deliberate practice, achievement, intervention

Supplemental materials: http://dx.doi.org/10.1037/pspp0000074.supp

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Deliberate practice—a challenging yet highly effective form of practice—leads to world-class eminence across domains. It helps turn chess players into grandmasters, violinists into virtuosos, and gymnasts into Olympians (Ericsson, 2006, 2008). Even critics, who point to other relevant determinants of skill improvement, including talent, acknowledge deliberate practice is an "unquestionably important" predictor of success (Macnamara, Hambrick, & Oswald, 2014, p. 1615).

Lauren Eskreis-Winkler, Psychology Department, University of Pennsylvania; Elizabeth P. Shulman, Psychology Department, Brock University; Victoria Young, Psychology Department, University of Pennsylvania; Eli Tsukayama, Psychology Department, University of Southern California; Steven M. Brunwasser, School of Medicine, Vanderbilt University; Angela L. Duckworth, Psychology Department, University of Pennsylvania.

This research was made possible by the National Institute on Aging (Grants K01-AG033182-02 and R24-AG048081-01), the Character Lab, the Gates Foundation, the Robert Wood Johnson Foundation, the Spencer Foundation, and the Templeton Foundation. The fifth author was supported in part from an NIMH training grant (T32-MH18921) during completion of this work. We are especially grateful to Rich Nelson, Vice President of ACCO brands, for his generous donation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. We are grateful to the students, teachers, and administrators who made this research possible. We thank Jonathan Baron, Philip Tetlock, Martin E. P. Seligman, Chris Hulleman, Carl Wieman, Scott Barry Kaufman, Brian Galla, Annie Park, and Yuchen Lin for their insightful feedback on the article. The first author extends special thanks to Jeffrey D. Winkler, and to her personal deliberate practice coach, Beverly D. Eskreis, a model of excellence in all things. Last but not least, thank you to the extraordinary, magnificent, inimitable Ari B. Lustig.

Correspondence concerning this article should be addressed to Lauren Eskreis-Winkler, Department of Psychology, University of Pennsylvania, 3701 Market Street, Suite 232, Philadelphia, PA 19104. E-mail: eskreisl@sas.upenn.edu

Journal of Personality and Social Psychology, 2016, Vol. 111, No. 5, 728–744
© 2016 American Psychological Association
0022-3514/16/$12.00 http://dx.doi.org/10.1037/pspp0000074

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

If deliberate practice is so effective at improving performance, why isn't it more commonly taught? Can non-experts be motivated to practice deliberately? Across four longitudinal, randomized-controlled, field experiments, we examined whether interventions that targeted expectancies and values, two established antecedents of ordinary effortful behaviors, could motivate deliberate practice among nonexperts and improve achievement over time. Our intervention built on foundational research which documents that brief, psychologically "wise" interventions—theoretically informed interventions that "take as their primary target a change in a specific psychological process" (Walton, 2014, p. 73)—can generate surprisingly large and durable effects (Garcia & Cohen, 2012; Walton, 2014; Wilson, 2011; Yeager & Walton, 2011).

Deliberate Practice and Achievement

In deliberate practice, one engages in tasks "that are initially outside [one's] current realm of reliable performance, yet can be mastered within hours of practice by concentrating on critical aspects and by gradually refining performance through repetitions after feedback" (Ericsson, 2006, p. 692). Deliberate practice has four main components: (a) the individual must have a well-defined goal for improving a specific aspect of performance, (b) the level of challenge should just exceed the individual's skill, (c) the individual receives immediate feedback, and (d) the individual repetitively focuses on the correction of error. Deliberate practice is associated with high levels of accomplishment in ballet, computer programming, aviation, firefighting, music, medicine, sports, and sales, among other fields (Ericsson, 2006, 2008). Most pertinent to the current investigation, undergraduates who do deliberate practice when they study earn higher grade point averages (GPAs; Plant, Ericsson, Hill, & Asberg, 2005).

There have been only a few attempts to increase deliberate practice via intervention. In these studies, conducted almost exclusively in the medical domain, clinicians are randomized to learn a medical procedure via deliberate practice (e.g., computerized simulators that offer immediate feedback and opportunities for repetitive practice) or training as usual (e.g., clinical rounds; Kessler et al., 2011). A recent meta-analysis (total $N =$ 633) confirms that deliberate practice leads to faster acquisition of medical skills than traditional training, $d = .71$, $p < .001$ (McGaghie et al., 2011). To our knowledge, only one study has assessed a deliberate practice intervention in a nonmedical, academic setting: Undergraduate physics students randomized to learn about electromagnetism via deliberate practice (e.g., they were given preclass reading quizzes, in-class clicker questions, and other activities designed to identify weaknesses and provide feedback) scored higher on a follow-up assessment than peers who received traditional lectures (Deslauriers, Schelew, & Wieman, 2011).

These past interventions demonstrate that when curricula are reengineered to require deliberate practice, faster learning ensues. In the present investigation, we address a novel question: Can deliberate practice be *motivated*? Can an intervention spur individuals to freely initiate and sustain deliberate practice? Whereas correlational, longitudinal research documents that individuals who engage in more deliberate practice achieve to higher levels than their peers (Ericsson, 2006), and experimental research demonstrates that redesigning a course can promote deliberate practice, to date, next to nothing is known about whether deliberate practice can be motivated.

Motivating Deliberate Practice

Deliberate practice, is ". . . [not] for the faint of heart" (Ericsson, Prietula, & Cokely, 2007). According to dancer Martha Graham (1998), practice involves "times of complete frustration . . . daily small deaths" (p. 95). Across domains, most individuals report that deliberate practice is, if not deadly, effortful and unpleasant (Duckworth et al., 2011; Ericsson, 2006, 2007, 2009; Ericsson & Ward, 2007; Ericsson, Krampe, & Tesch-Romer, 1993). Performers in all domains, including sports, find that mental exhaustion limits how much deliberate practice they complete (Ericsson, 2002, 2003). Even seasoned experts sustain, on average, no more than four hrs of deliberate practice per day (Ericsson, 1996; Ericsson et al., 1993). Long before academic research on deliberate practice began, Bryan and Harter (1897) concluded that telegraph operators plateaued in their skill level because they would "not make the painful effort necessary to become experts" (p. 51).

What motivates some to undertake these "painful efforts?" Two related psychological antecedents of deliberate practice have been identified: grit (Duckworth et al., 2011) and harmonious passion (Bonneville-Roussy, Lavigne, & Vallerand, 2011; Vallerand et al., 2007; Vallerand et al., 2008). Grit is the tendency to sustain effort and interest toward a single pursuit over time (Duckworth et al., 2007). In a study of National Spelling Bee finalists, hours of deliberate practice fully mediated the prospective association between grit and final competition performance (Duckworth et al., 2011). Similarly, individuals with harmonious passion—those who have internalized cherished activities into their identities—do more deliberate practice, and this explains their higher levels of achievement in basketball, ballet, music, and other domains (Bonneville-Roussy, Lavigne, & Vallerand, 2011; Vallerand et al., 2007; Vallerand et al., 2008).

Building on this correlational research, we conducted an experimental test of a causal hypothesis: Could targeting established determinants of effortful behavior *motivate* deliberate practice? We moved away from expert, high-achieving populations to work with nonexpert students who, as a group, do not exhibit unusual motivation to invest effort in deliberate practice in any domain. In doing so, we addressed the clarion call of Ericsson and Charness (1994), who urged scientists to attain "A better understanding of . . . factors that motivate and sustain . . . deliberate practice" (p. 745).

An Expectancy-Value Approach to Intervention

To motivate deliberate practice, our intervention targeted expectancy and value, two key determinants of effortful behavior (Atkinson, 1957; Battle, 1965; Crandall, 1969; Eccles & Wigfield, 2002; Feather, 1982; Pintrich, 2003; Wigfield & Eccles, 2000). According to expectancy-value theory (EVT), expectancy is the extent to which people believe they will succeed. Value, which refers to the subjective value the individual attaches to success, varies positively with perceived benefits (e.g., attainment value, intrinsic value, utility value) and inversely with perceived costs (e.g., the opportunity cost, emotional cost, and/or required effort). Expectancy and value are reliably associated with effort expendi-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

ture and achievement among nonexperts (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000), particularly in academic settings (Eccles, Midgley, & Adler, 1984; Eccles, Wigfield, Harold, & Blumenfeld, 1993; Meece, Wigfield, & Eccles, 1990; Nagengast, Trautwein, Kelava, & Ludtke, 2013).

Interventions that target expectancy and/or value have been shown to improve academic performance. For example, growth mindset interventions, which teach students that the brain can grow with effort, improve end-of-year grades, particularly for lower-achievers (Aronson, Fried, & Good, 2002; Blackwell et al., 2007; Good, Aronson, & Inzlicht, 2003; for review, see Yeager & Dweck, 2012). Similarly, interventions that target perceived costs, a component of value in EVT, improve academic achievement. In one study involving sixth graders, students exposed to a 10-min intervention that encouraged them to reappraise learning costs (e.g., difficulty during learning) as normative outperformed controls on a follow-up reading comprehension task (Autin et al., 2012). Likewise, undergraduates taught to reappraise the anxiety experienced during high-stakes exams as a *good* sign scored higher on the Graduate Record Examination (GRE) several months later (Jamieson et al., 2010).

The present intervention paralleled past interventions by targeting expectancies and costs, yet placed a novel focus on the role of practice. For example, in addressing expectancy, the intervention taught: "Many people think talent is all that matters . . . actually scientific evidence suggests that *deliberate practice* is incredibly important to improvement and success." This message is distinct from the focal message of growth mindset interventions, which teaches students that if they invest effort, their brains will grow. We believed it would be important to instill this practice-specific expectancy given the prevalence of the opposing belief—that talent, not practice, determines success (Tsay & Banaji, 2011). The intervention also encouraged students to reappraise costs specific to practice: "If you are frustrated or confused *while practicing* . . . it can mean you are working on your weaknesses . . . when you practice and everything goes perfectly, it may feel good, but it's probably a sign that you're not challenging yourself." We thought it would be crucial to introduce this reappraisal because expending effort is inherently aversive and, all other things equal, avoided (Eisenberger, 1992; Gray, 2000; Kool, McGuire, Rosen, & Botvinick, 2010; Smith & Walker, 1993). Our focus on practice-related costs differentiated the present intervention from past interventions that have encouraged students to reappraise either learning as a whole (Autin et al., 2012) or the specific anxiety experienced during exams (Jamieson et al., 2010).

Intervening Wisely

While it may seem surprising to suggest that a brief intervention could affect academic achievement over time, psychologically "wise" interventions have been shown to produce surprisingly large and durable effects on effort expenditure and achievement, among other outcomes (Blackwell et al., 2007; Walton & Cohen, 2011; Aronson, Fried, & Good, 2002; Cohen et al., 2009; Hulleman & Harackiewicz, 2009; for reviews, see Garcia & Cohen, 2012; Yeager & Walton, 2011). How do they work? Wise interventions target key psychological processes that influence outcomes of interest (Lewin, 1952; Ross & Nisbett, 1991). Targeting such processes can generate initial benefits and, through recursive processes, perpetuate themselves long-term. For example, a student who comes to believe more strongly in the importance of effort might study harder and perform better in school. Improved performance may then reinforce the belief that effort matters, jumpstarting the cycle anew. In the current investigation we designed a brief, wise intervention in the hopes of changing deliberate practice habits and achievement over time.

Helping the Least Expert

We expected the intervention would improve beliefs and deliberate practice behavior for all students, but preferentially boost the achievement of lower-performers—the least expert of the nonexperts. Wise interventions often differentially improve the performance of lower-achieving students (Cohen et al., 2009; Hulleman & Harackiewicz, 2009; Paunesku et al., 2015; Wilson & Linville, 1982; Yeager et al., 2014a). This may be due to ceiling effects: Students already earning A's cannot improve their grades, but students earning D's, C's, or B's have room to grow. Because most students in nonexpert settings do less deliberate practice than they are capable of, we believe almost all students stand to benefit from learning about deliberate practice. However, given that there is an artificial ceiling on measures of achievement in the academic context (e.g., GPA), we expected the intervention would preferentially raise the grades of lower-performing students.

The Current Investigation

Across four longitudinal, randomized-controlled, field experiments, we examined the effect of a wise intervention on expectancy-value beliefs, deliberate practice, and achievement. As a preliminary (Study 1), we validated a novel behavioral measure of deliberate practice and examined the relationship between deliberate practice and achievement in the nonexpert middle school setting. Using this task, we also confirmed associations between deliberate practice and expectancy-value beliefs. Next, we intervened. We tested the effect of a wise, 25-min intervention on middle school math performance (Study 2) and college grades (Study 3). Finally, we developed an expanded 50-min intervention for middle school students—sixth (Study 4) and seventh (Study 5) graders—and evaluated its effects on academic achievement, self-reported beliefs, and deliberate practice behavior, measured at one-month (Study 4) and four-month (Study 5) follow-up. Whereas in Studies 2 and 3 the control condition instructed students in study skills, in Studies 4 and 5 it discussed interests and achievement. This combination of studies allowed us to test the efficacy of the intervention across a range of students and a range of outcomes.

Because each wise intervention we administered was both informational (e.g., an introduction to the tenets of deliberate practice) and motivational (e.g., content more explicitly aimed at changing self-relevant expectancies and values associated with practice), in Study 5, we introduced an additional condition in which students received information about deliberate practice without an accompanying motivational lesson. With this hybrid condition, we parsed the intervention's informational effects from its motivational ones to isolate active ingredients.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Study 1

In Study 1, we assessed the importance of deliberate practice to nonexpert achievement and examined the association between expectancy-value beliefs and deliberate practice in the middle school setting. Diary measures and retrospective self-report measures are traditionally used to measure deliberate practice (Ericsson et al., 1993; for review, see Ericsson, 2006); however, we doubted whether average middle school students would be able to reliably distinguish between deliberate practice and less effective forms of practice, a distinction existing self-report measures rely on. Indeed, our intervention approach is premised on the assumption that deliberate practice is not widely understood among nonexperts. Therefore, we developed a novel, objective task measure of deliberate practice.

In the deliberate practice task (DPT), students work through a series of challenging online math problems in an online browser during a 45-min class period. At any point, students may switch away from the math task to YouTube, Facebook, Instagram, or other recreational websites by opening new tabs. The math task captures several important features of deliberate practice: It keeps students in a zone of challenge by leveling them up to harder problems when they answer easier ones; it offers students the opportunity to view immediate feedback (e.g., worked out answers with explanations); finally, it encourages repetition by presenting iterations of similar problems until mastery is achieved. Because the DPT captures key features of deliberate practice, deliberate practice was operationalized as the amount of time students chose to spend on-task (*time focused*), as opposed to off-task, during a 45-min class period.

We first examined whether the DPT demonstrated convergent validity with Big Five conscientiousness and its facets, grit and self-control. We also tested whether the DPT demonstrated discriminant validity from dissimilar constructs, such as Big Five extraversion, agreeableness, openness, and emotional stability. Next, we examined the incremental predictive validity of the DPT for GPA, an objective measure of academic achievement. Finally, we assessed the relation between the DPT and two hypothesized psychological antecedents of deliberate practice: expectancies and values.

Method

Participants and procedure. Participants were sixth (n = 542) and seventh (n = 417) graders attending two school districts in the U.S. Approximately 68% were White, 15.3% Asian, 8.5% Hispanic, 8.2% Black; 51.1% were female. Seven percent of students qualified for free or reduced price meals, an indication of low household income.

Students completed all activities on school computers over three consecutive days. On Days 1 and 2 students completed self-report measures; on Day 3 students completed the DPT. During all research sessions, students wore ear buds and each computer was covered with a privacy screen—a plastic sheet that lessens visibility to anyone not seated directly in front of the computer.

Because some students did not complete some assessments, degrees of freedom varied by approximately 8% across analyses. Missing data were due to students opting out, technological malfunctions (e.g., computer crashing), and absences. In addition, we excluded DPT data from participants who spent less than five min signed into the task (approximately 2% of the sample). Short task time may have resulted from early dismissal, students accidentally closing the browser, students intentionally closing the browser, a technological glitch, or other causes. Pairwise deletion was used across analyses. In separate models (not shown), listwise deletion yielded virtually identical results. For more information on the final sample, see the online supplement.

Measures.

Deliberate practice task (DPT). The DPT included standard textbook math problems (e.g., $6x = 10$) at nine skill levels. Math problems and solutions were taken from Khan Academy, a free, educational web site (www.khanacademy.org). The task was tested with several students beforehand to ensure math content was age-appropriate. The task started all students with math problems at the easiest level, Level 1. After each set of five problems, students received their score (number correct out of five) and had the opportunity to view worked-out solutions to the problems they had completed. Students who answered five consecutive problems correctly were advanced to the next level. The aim of this leveling up feature was to keep students in a zone of constant challenge.

Before beginning the math task, students were told to treat the task like a regular school assignment. They were also told, "Because today's math lesson is hard, you are allowed to do something you're not usually allowed to do in school: take breaks on the internet." In coordination with school administrators, recreational websites (e.g., YouTube, Facebook, Instagram) were unblocked for the entire 45-min period, allowing students to open additional browser tabs and visit sites of their choosing.

Time focused. The amount of time students focused on math problems was measured using an embedded timer invisible to the user. This timer monitored the number of seconds students spent on the tab with math problems versus other tabs. Because the task approximated key features of deliberate practice (e.g., it leveled students up, it provided them with the opportunity to view feedback, it allowed for repetition), we operationalized deliberate practice as the amount of time students focused on the math task during the 45-min period.

Total task time. This was the total number of seconds that students were signed into the task during the 45-min class period, regardless of whether they were working on math (time focused) or surfing the internet.

Self-report questionnaires. Students rated all questionnaire items on a 5-point scale from 1 = *not at all like me* to 5 = *very much like me,* except where otherwise noted. Items for each scale were averaged to create a composite. Higher values corresponded to greater endorsement of the specified construct. We included two types of expectancy-value scales: (a) practice-specific scales (deliberate practice beliefs, frustration tolerance during practice), which narrowly assessed an individual's expectancies and values related to practice; and (b) nonpractice-specific scales, which assessed expectancies and values that were not directly related to practice (e.g., self-efficacy, locus of control, distress tolerance). For descriptive statistics, see Table 1.

Big Five personality. We adapted items from the Big Five Inventory to improve reading comprehension (John & Srivastava, 1999). We used four items per trait to assess Conscientiousness (e.g., "I am organized and neat;" α = .85), Agreeableness (e.g., "I am considerate and kind to almost everyone;" α = .73), Emotional

Table 1
Descriptive Statistics and Partial Correlations (pr) With Time Focused for Study 1

	M	*SD*	*n*	*pr*
Deliberate practice task				
Time focused (seconds)	1,127	378	885	—
Total task time (seconds)	1,448	305	885	—
Big Five Personality				
Big Five Conscientiousness	3.91	.77	831	$.22^{***}$
Grit	3.45	.66	843	$.21^{***}$
Self-control	3.31	.69	841	$.20^{***}$
Big Five Agreeableness	4.08	.65	832	$.11^{**}$
Big Five Emotional Stability	2.86	1.01	832	.06
Big Five Openness	3.98	.71	833	−.02
Big Five Extraversion	3.89	.69	835	$-.09^{**}$
Academic performance				
GPA	88.39	7.80	885	$.19^{***}$
Expectancy-value beliefs				
Practice-specific				
Deliberate practice beliefs	77.09	13.82	844	$.14^{**}$
Frustration tolerance	3.67	.68	844	$.15^{***}$
Nonpractice-specific				
Growth mindset	3.60	.86	844	.03
Locus of control	3.97	.57	844	$.11^{**}$
Distress tolerance	2.56	1.05	844	−.06

Note. All partial correlations control for school site, gender, ethnicity, standardized math test achievement, and total task time. Partial correlations with Big Five personality scales appear in descending order of magnitude.
** $p < .01$. *** $p < .001$.

Stability (e.g., "I get stressed out easily;" reverse scored; α = .76), Openness to Experience (e.g., "I am curious. I am very interested in learning new things;" α = .65), and Extraversion (e.g., "I am outgoing, I like to meet new people;" α = .64).

Grit. Students completed the eight-item (e.g., "I am a hard worker") Short Grit Scale (Duckworth & Quinn, 2009; α = .78).

Self-control. Students completed the 13-item Brief Self-Control Scale (e.g., "People would say that I have iron self-discipline;" Tangney, Baumeister, & Boone, 2004; α = .85).

Practice-specific expectancy and value. To assess practice-specific expectancies, students completed a measure of *deliberate practice beliefs*. In this measure, they were asked to "advise another student on how to succeed in school" by rating each of the five main components of deliberate practice (e.g., focus on weaknesses, concentrate 100%) using sliders that ranged from 0 = *not at all helpful* to 100 = *very helpful*. Due to time constraints, this measure was only administered in one of the two school districts. To assess costs, one component of value in EVT, students answered a three-item measure of frustration tolerance (e.g., "If you are frustrated and struggling when practicing, it is a sign that you are challenging yourself and improving"), which measured the degree to which participants endorsed frustration *during practice* as a sign of learning (α = .44). Removing the one nonreverse-scored item from the frustration tolerance scale improved its alpha reliability (α = .63). Rerunning analyses with this one item removed led to no significant changes in the reported results, so the full three-item scale was used in all reported analyses.

Nonpractice-specific expectancy and value. To assess nonpractice-specific expectancy, participants completed two standard, less specific measures of expectancy: the three-item (e.g., "You have a certain amount of intelligence and you cannot change it;" reverse scored) Growth Mindset Scale (Dweck, Chiu, & Hong, 1995; α = .72) and the four-item (e.g., "If your teacher gives you a bad grade it's not your fault;" reverse scored) Locus of Control Scale (Wellborn, Connell, & Skinner, 1989; α = .46). To assess nonpractice-specific costs, participants completed three items (e.g., "Feeling upset is unbearable to me;" reverse scored) that assessed distress tolerance (α = .82).

Grade-point average (GPA), standardized math achievement test scores, and demographics. From school records, we obtained first-quarter GPA for all core subjects (English, social studies, science, math), standardized math achievement test scores from the prior year, gender, and ethnicity. Because measures of achievement were on different scales at different school sites, these measures were standardized within schools prior to analyses.

Results and Discussion

Preliminary analyses. On average, students correctly solved only 22% of the math problem sets they attempted, suggesting that the task successfully kept students in a zone of challenge. Controlling for demographics (e.g., gender, ethnicity), school affiliation, and total survey time, students with higher standardized math achievement scores were more focused ($pr = .13$, $p < .001$; for additional descriptives, see Table 1).

Next, we conducted a series of partial correlations to examine convergent, discriminant, and incremental predictive validity, as well as the association between deliberate practice and expectancy-value beliefs. Across analyses, we controlled for standardized math achievement test scores from the prior year, ethnicity, gender, school affiliation, and total task time. Results were not moderated by these covariates in this study or subsequent studies.

Convergent and discriminant validity. Time focused was associated with Big Five Conscientiousness ($pr = .22$, $p < .001$), grit ($pr = .21$, $p < .001$), self-control ($pr = .20$, $p < .001$), Agreeableness ($pr = .11$, $p = .002$), and inversely, with Extraversion ($pr = -.09$, $p = .01$; see Table 1). Time focused was unrelated to Big Five Openness and Emotional Stability ($prs = -.02, .06$, $ps = .55, .12$). Because Big Five factors were intercorrelated (average $r = .19$), we fit a simultaneous regression model with all the Big Five factors predicting time focused. In this model, only Conscientiousness ($\beta = .21$, $p < .001$) and Extraversion ($\beta = -.14$, $p < .001$) continued to predict unique variance. The negative correlation between Extraversion and time focused parallels past research in which extraverted students earn lower grades (Poropat, 2009).

Given the conceptual similarity between grit and Conscientiousness, and the high correlation between these constructs in the present data set, $r = .73$, $p < .001$, we fit a simultaneous regression model to see whether grit predicted incremental variance in the tendency to stay focused over and above Conscientiousness. It did (grit: $\beta = .09$, $p = .02$; Conscientiousness: $\beta = .10$, $p = .01$).

Criterion-related validity. Students who spent more time focused earned higher GPAs ($pr = .18$, $p < .001$). This association did not change when math grades were excluded from the GPA calculation ($pr = .18$, $p < .001$), confirming that the DPT captured practice behaviors associated with achievement in general, not just math achievement.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Expectancy-value beliefs. Time focused correlated with practice-specific expectancies (deliberate practice beliefs, $pr = .14$, $p < .01$), and values (specifically, costs, one component of value in EVT; frustration tolerance, $pr = .15$, $p < .001$), but was not consistently associated with nonpractice-specific expectancies (growth mindset, $pr = .03$, $p = .36$; locus of control, $pr = .11$, $p < .01$) or values (again, we assessed the cost component; distress tolerance, $pr = -.06$, $p = .08$).

Study 2

In Study 1, we validated a novel behavioral measure of deliberate practice, established the importance of deliberate practice to middle school achievement, and uncovered associations between expectancy-value beliefs—particularly those that were practice-specific—and deliberate practice. Building on these encouraging findings, in Study 2, we developed and administered a wise, 25-min intervention to fifth and sixth graders. Students were randomized to either a treatment condition that aimed to motivate deliberate practice or an active control condition that taught standard study skills. Approximately 1 week later, all participants were instructed to master math concepts in Khan Academy (a free educational web site). We evaluated whether this brief intervention improved math achievement.

Method

Participants. Students were fifth ($n = 98$) and sixth graders ($n = 111$) at four schools across the U.S. Approximately 53% of participants were White, 23.9% Black, 11.0% Hispanic, 9.1% Asian, and 5.3% were of other ethnicities; 44.0% were female.

Procedure.

Overview. In the spring, participating schools administered a 25-min online intervention to students. Because the intervention contained audio and video, students wore ear buds during all research sessions. In this study and all subsequent studies, an algorithm in Qualtrics, the online platform used to host the intervention, randomly assigned students to condition. Immediately prior to the intervention, participants completed a short math pretest.

During the second research session and all subsequent sessions, students were instructed to master multiplication and division content in Khan Academy. On average, the second research session took place five school days after the intervention. Schools allotted varying numbers of class periods to Khan Academy (anywhere from five to nine class periods on consecutive school days) based on computer availability.

In this study and all subsequent studies, teachers were completely blind to condition assignment and intervention content. Teachers were not shown the intervention modules until the study was over, and were not informed in advance about the specific content being taught to students.

Intervention content. Treatment and control modules were made as parallel as possible. They were matched for amount of text, video length, number of images, and number of activities. Moreover, across treatment and control conditions, we employed two psychological tactics commonly used to make wise interventions "stickier:" descriptive norms and the saying-is-believing effect (Aronson, 1999; Aronson, Fried, & Good, 2002; Walton & Cohen, 2011; Yeager et al., 2014a; Yeager et al., 2014b). To incorporate descriptive norms, we showed participants anonymous quotes in which other students (who had participated in focus groups) described their practice habits and preferences. These quotes were edited by our research team for clarity and brevity. Sharing such quotes normalized the behaviors the intervention encouraged (Cialdini, Reno, & Kallgren, 1990; Goldstein, Cialdini, & Griskevicius, 2008; see Cialdini, 2003). Participants also completed "saying-is-believing" exercises, in which they wrote letters to other students advocating the importance of what they had learned. These exercises capitalize on the psychological insight that one of the most effective ways to persuade a participant of a message is to have the participant advocate the message to others.

Treatment condition. The treatment condition consisted of two parts. In Part 1, students learned the tenets of deliberate practice: (a) focus on weaknesses, (b) get feedback, (c) concentrate 100%, and (d) repeat until mastery. Didactic slides were interspersed with activity prompts, an illustrative video, and a letter-writing exercise. The aim of Part 1 was to teach students about deliberate practice and how it differs from less effective forms of practice.

Part 2 was motivational. It targeted expectancies and values, two psychological variables highlighted by EVT. To address expectancies, the module taught that talent and effort both contribute to success, but that the relative importance of effort—especially effort invested in deliberate practice—is often underestimated. To address costs, a component of value in EVT, the module encouraged students to interpret frustration and confusion as positive signs that one is engaging in optimal practice activities. Thus, the intervention targeted practice-specific expectancies and costs in an attempt to motivate deliberate practice. For a fuller description of the procedure used to develop the intervention, the full intervention text, as well as a table showing how specific quotes from the intervention map onto each theoretical concept, refer to the online supplement.

As noted above, each lesson ended with a saying-is-believing exercise in which the participant wrote a letter to another student endorsing deliberate practice. For example, one treatment participant wrote:

> Deliberate practice is studying weaknesses or something that you are struggling at. How I applied this is by doing flashcards on things I do not know. At first I was using regular practice and everyone seemed so much better than me but when I started using deliberate practice my grades went up.

Control condition. Control participants were taught standard study advice. To rule out the possibility that any discussion of studying or practice would be as effective as teaching students about deliberate practice, controls learned how to manage their time (e.g., use a calendar or planner) and were given practical strategies for remembering information (e.g., mnemonics). Parallel to the treatment module, the control module featured short activity prompts and an illustrative video.

Like the treatment condition, the control condition featured descriptive norms, and ended with students writing a letter to another student, endorsing what they had learned. For example, one control participant wrote:

> I have learned many new tricks for studying. I thought that the mnemonics were really cool. You can use the things you already know to make things or work easier. Mnemonics keep you hooked on them books. Next time I think I will use this skill to help myself to get through a hard problem. Studying is the key to success.

Measures.

Khan Academy performance, math pretest, and basic demographics. Performance was quantified as the total number of points earned in Khan Academy. Students could earn up to 100 points in each of 19 multiplication and division topic areas, for a maximum of 1,900 points.

The math pretest had several multiplication and division problems taken from Khan Academy. Khan Academy performance and pretest scores were standardized by site before being combined into a single variable. Before the math pretest, students self-reported gender and ethnicity.

Interest and engagement. At the end of the intervention, students rated how interesting and engaging they found the module. Students rated two adjectives ("interesting" and "engaging") using 100-point slider scales. The two values, $r = .72$, $p < .001$, were averaged into a single composite.

Results and Discussion

Analytic plan for Studies 2 through 5. We analyzed effects using ordinary least squares regression. Regression models controlled for basic demographics (e.g., gender, ethnicity, school site) and available measures of baseline achievement (e.g., pretest scores, prior GPA, standardized test scores from the prior year). Because measures of achievement were on different scales at different school sites, achievement measures were standardized by site prior to analyses.

Because we expected the intervention to preferentially help lower-achievers, we tested for main effects and Condition × Prior Achievement interactions. We probed all significant interactions using the Johnson-Neyman method (Hayes, 2013; Johnson & Neyman, 1936) to identify regions of significance. We ran analyses using R statistical software (available at http://www.r-project.org) as well as the SPSS PROCESS macro (Hayes, 2013).

Preliminary analyses. Treatment and control groups did not differ on baseline achievement or demographics, with the exception of baseline ethnicity (for more details, see online supplement). There was no between-condition difference in how interesting and engaging students found the modules, $t(187) = -1.14$, $p = .26$, $d = .16$, 95% CI [−.12, .45].

Khan Academy performance. Although the main effect of the intervention was not significant ($M_{treatment} = 0.14$, $M_{control} = 0.05$),[1] $t(198) = 0.64$, $p = .52$, $d = .09$,[2] 95% CI [−.18, .36], the intervention improved online math performance for lower-achievers, Condition × Math Pretest interaction, $b = -0.37$, $SE = .14$, $t(197) = -2.65$, $p = .01$ (see Figure 1). A follow-up Johnson-Neyman analysis (Hayes, 2013; Johnson & Neyman, 1936) revealed that the effect of treatment was significantly positive among lower performing students (< -0.76 *SD* below average on the pretest) and significantly negative among high-performing students (> 1.78 *SD* above average on the pretest). Although unexpected, the negative effect among high-performing students may have been due to the fact that the treatment condition advised students to focus on areas of challenge and weakness. For students who had already mastered the assigned math content (as evidenced by high pretest scores), Khan Academy performance did not constitute the sort of challenge the intervention encouraged them to confront. The intervention may have discouraged these students from working on problems that were too easy for them.

Study 3

In Study 2, a 25-min intervention improved math performance for lower performing middle school students. We next tested whether this effect would generalize to older students and an objective measure of academic achievement: college GPA.

Method

Participants. Participants were undergraduates at a liberal arts college ($n = 60$) and a research university ($n = 60$). Students were predominantly White (71.7%), female (69.2%), and in their first 2 years of college (74.2%).

Procedure.

Overview. At the liberal arts college, students were recruited through an e-mail sent to the entire student body (approximately 850 students) 3 weeks prior to finals. Participants were entered in a raffle to win a $100 Amazon gift certificate. A total of 63 students (approximately 7%) participated. Three students with no spring GPA on record were dropped from analyses. At the research university, the intervention was advertised on the psychology department's website. Sixty students in a large introductory psychology course (14%) participated for course credit. Participation was open to students for 2 weeks prior to the start of reading days.

Intervention content. The intervention was nearly identical to the intervention administered in Study 2. Language was slightly modified to ensure it was age-appropriate, but content remained unchanged.

Measures.

Academic achievement and basic demographics. Academic achievement was evaluated differently by site. The liberal arts college provided overall fall and spring semester GPAs, on 4-point scales. The research university released midterm and final grades for Psychology 001, on 100-point scales. Fall GPA (at the liberal arts college) and midterm scores (at the research university) were used as measures of prior achievement. Gender and ethnicity were also obtained from official school records.

Interest and engagement. We assessed interest and engagement using the same questions used in Study 2. As in Study 2, responses, $r = .71$, $p < .001$ were averaged into a single composite.

Results and Discussion

Preliminary analyses. Randomization was successful. Treatment and control participants did not differ on any baseline variables (for details, see online supplement). Nor did they differ in

[1] These means and all subsequent reported means are covariate-adjusted. They were calculated using the *lsmeans* package in R.

[2] For all intervention effects, Cohen's *d* is calculated by dividing the unstandardized beta coefficient by the pooled standard deviation (see Feingold, 2015, Equation 3).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

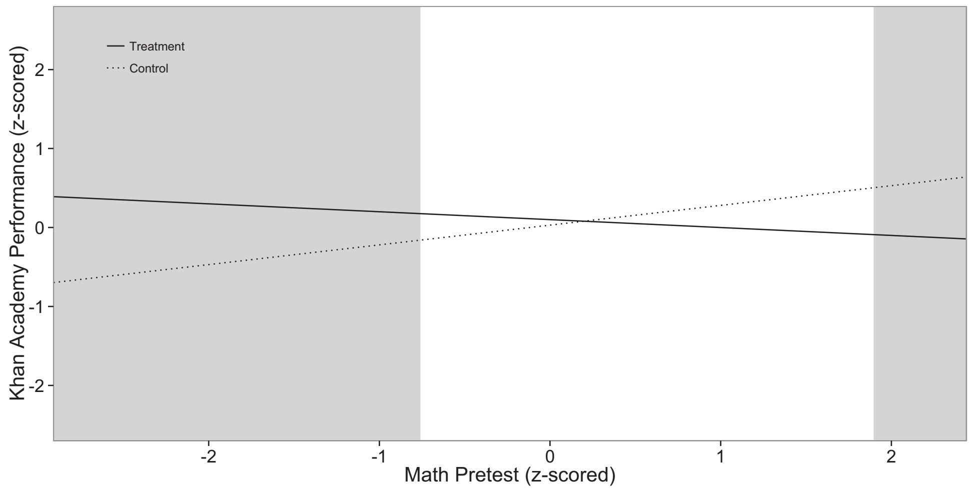

Figure 1. In Study 2, the treatment condition improved math performance in Khan Academy for lower achievers, and the control condition improved math performance in Khan Academy for higher achievers. Shaded areas indicate regions of significance.

how interesting and engaging they found the modules, $t(110) = -1.31$, $p = .19$, $d = .25$, 95% CI [−.12, .62].

Academic achievement. The intervention had a main effect on end-of-semester grades ($M_{\text{treatment}} = 0.10$, $M_{\text{control}} = -0.19$), $t(111) = 2.03$, $p = .04$, $d = .38$, 95% CI [.01, .74]. This main effect was driven by lower-achievers: Condition × Prior Achievement interaction, $b = -.37$, $SE = .15$, $t(110) = -2.51$, $p = .01$ (see Figure 2). A follow-up Johnson-Neyman analysis (Hayes, 2013; Johnson & Neyman, 1936) revealed that the effect of treatment was significantly positive among lower-performing students (<0.08 *SD* above average on prior achievement).

Study 4

In Studies 2 and 3, a brief psychological intervention improved academic performance among lower-achieving students in middle school (Study 2) and college (Study 3). In Study 4, we examined the psychological and behavioral processes the intervention activated. We used questionnaires to evaluate self-reported expectancies and values, and the DPT (developed and validated in Study 1) to measure deliberate practice behavior. Expectancies, values, and deliberate practice were measured once preintervention and twice postintervention. As in Study 1, we measured practice-specific as well as nonpractice-specific expectancies and values.

In addition, we introduced an expanded intervention and a new control condition. Feedback from focus groups revealed that students were most engaged by interactive material. We therefore created a longer, 50-min intervention that expanded on both Part 1 and Part 2 of the earlier intervention, by including more activities. In addition, to rule out the possibility that merely mentioning motivation and high achievement drove the treatment effect in

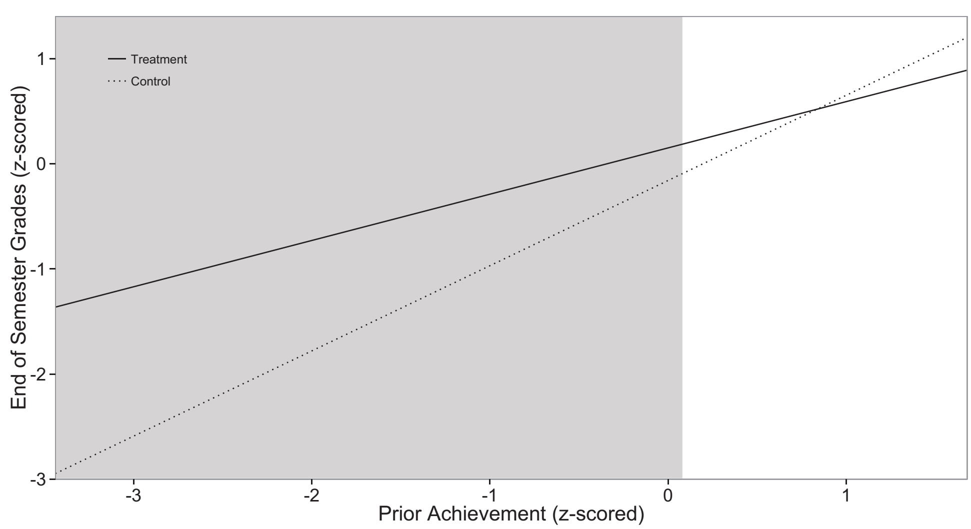

Figure 2. In Study 3, the intervention raised end-of-semester academic achievement for all undergraduates, and especially among lower achievers. Shaded areas indicate regions of significance.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Studies 2 and 3, control participants received a new control module that discussed motivation and achievement. In particular, students learned how deep interests develop and how deep interests might promote achievement in their own lives.

Finally, in Study 4, we tested the durability of the intervention's effect on achievement. Whereas in Studies 2 and 3 we measured academic achievement several weeks later, in Study 4 we measured the intervention's impact on achievement over an entire academic quarter (approximately 11 weeks later). Our timeline was informed by past studies which document that brief, wise interventions can generate positive effects on achievement that last months, and even years (Aronson, Fried, & Good, 2002; Blackwell et al., 2007; Cohen et al., 2009; Hulleman & Harackiewicz, 2009; Walton & Cohen, 2011; for reviews, see Garcia & Cohen, 2012; Yeager & Walton, 2011). One theory is that wise interventions generate long-term effects by targeting key psychological and behavioral processes in recursive settings, that is, those in which small changes build over time (Walton, 2014; Yeager & Walton, 2011). Schools are inherently recursive environments: later lessons build on earlier content. Drawing on this foundational research, we tested the long-term effects of a brief intervention that targeted practice-specific beliefs (e.g., expectancies, costs) and deliberate practice.

Method

Participants. We invited sixth graders in a public school district in the United States to participate. Approximately 78% of students (N = 427) across three middle schools participated. Remaining students either opted out, experienced technological difficulties (e.g., computer crashed), or were absent. Participants did not differ from nonparticipants on any baseline variables (e.g., incoming GPA, gender, ethnicity). For further details, see the online supplement.

Approximately 65% percent of participants were White, 19.4% Asian, 7.5% Hispanic, 4.9% Black, and 2.8% were of other ethnicities; 48.7% were female. Approximately 7% of students qualified for free or reduced price meals.

Procedure.

Overview. First, students completed a battery of self-report scales and the DPT during three class periods on three consecutive days. Ten weeks later, at the end of the third quarter, students completed the intervention during two class periods on two consecutive days. One day after the intervention (Follow-up 1) and again 1 month later (Follow-up 2), students completed the same battery of measures they had taken preintervention (i.e., self-report scales and the DPT). During all research sessions, students wore ear buds and privacy screens were attached to all computers. GPA was obtained at one time point preintervention (second quarter GPA) and at one time point postintervention (fourth quarter GPA).

Intervention procedures and content. Treatment and control modules employed the same persuasive techniques (descriptive norms and the saying-is-believing effect) described in Study 2. Modules were carefully matched to contain the same-length text, same-length videos, and parallel activities. Across both conditions, didactic slides were interspersed with interactive multiple-choice questions, short-answer prompts, a self-quiz, and an illustrative video (for sample screenshot, see Figure 3).

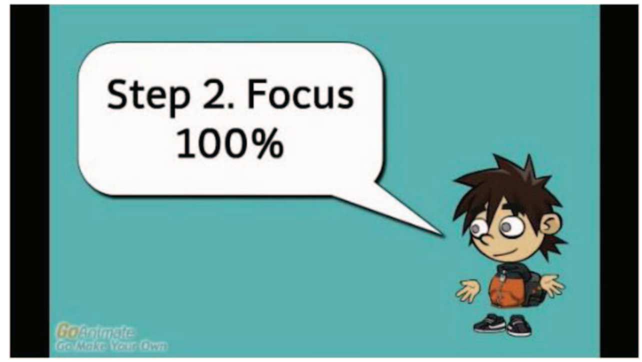

Figure 3. Screenshot of the intervention. See the online article for the color version of this figure.

Treatment condition. As in Studies 2 and 3, the intervention had two parts. In Part 1 (25 min), students were taught the tenets of deliberate practice (refer to the description in Study 2 for more details). During focus groups, students preferred the term "deep practice" to "deliberate practice," so "deep practice" was substituted. The aim of Part 2 (25 min) was to motivate students to use the deliberate practice techniques they learned about in Part 1. As in the earlier intervention (Studies 2 and 3), Part 2 targeted expectancies and costs with the aim of encouraging students to reappraise practice-specific expectancies and values. Refer to the online supplement to see the full intervention text as well as a table which shows how specific quotes from the intervention map onto each theoretical concept.

Control condition. In Part 1, students learned about deep interests and high achievement. They learned what it looks and feels like to be motivated, how to increase motivation, and how to work around the obstacles that stand in the way of motivation and high achievement (e.g., students were cautioned against relying on what their friends do and were told that instead, they should follow their own interests). The active motivational beliefs targeted in the treatment condition—expectancies and values—were not touched upon.

Building on Part 1, Part 2 discussed the link between deep interests, motivation, and success as well as what being motivated feels like. In exercises and activities that paralleled those in the treatment condition, students were encouraged to apply what they had learned to their own lives.

Measures.

Self-report questionnaires. Recall that the wording of the intervention aimed to increase the expectancy that *deliberate practice* would lead to success and reduce the cost associated with *deliberate practice*. Because the intervention addressed expectancy and cost specifically as they relate to deliberate practice, we included practice-specific measures of expectancy (deliberate practice beliefs at preintervention, α = .45; at Follow-up 1, α = .60; at Follow-up 2, α = .67) and cost (frustration tolerance at preintervention, α = .38; at Follow-up 1, α = .63; at Follow-up 2, α = .56), as well as nonpractice-specific measures of expectancy (growth mindset at preintervention, α = .75; at Follow-up 1, α = .80; at Follow-up 2, α = .84; locus of control at preintervention, α = .45; at Follow-up 1, α = .68; at Follow-up 2, α = .72) and cost (distress tolerance at preintervention, α = .81; at Follow-up 1,

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

$\alpha = .86$; at Follow-up 2, $\alpha = .90$). For full details on these measures, see Study 1.

Removing the one item from the frustration tolerance scale improved its alpha reliability preintervention ($\alpha = .54$). However, because rerunning analyses with this one item removed led to no significant changes in the reported results, the full three-item scale was used in all reported analyses.

DPT, time focused, and total task time. Participants completed the same task described in Study 1.

GPA, standardized test scores, and basic demographics. GPA, standardized achievement test scores from the prior year, gender, and ethnicity were obtained from school records. Because the intervention was administered in the last several weeks of the third quarter, we assessed the intervention's effect on fourth quarter GPA, controlling for GPA from the prior quarter (second quarter).

Interest. Students rated how interesting they found each module on a 100-point slider scale. We averaged ratings for the two modules, $r = .77$, $p < .001$, to create a single measure of student interest.

Results and Discussion

Analytic plan for Studies 4 and 5. As in Studies 2 and 3, effects on GPA were analyzed using ordinary least squares regression (see Study 2 for further details on the analysis plan for regression). Gender, ethnicity, school site, standardized achievement test scores, and second quarter GPA were included as covariates in all models. Following precedent (e.g., Yeager et al., 2014b), we controlled for standardized achievement test scores and prior GPA to comprehensively account for incoming differences in student achievement. As in Studies 2 and 3, across Studies 4 and 5, achievement measures were standardized by site prior to analyses.

In contrast to GPA, beliefs and behaviors were assessed at multiple time points (once before and twice after the intervention). Changes in beliefs and behaviors were analyzed using linear mixed models in R statistical software (*nlme* package, v. 3.1). Missing data were considered missing at random (MAR). Parameters were estimated using restricted information maximum likelihood (REML; Singer & Willett, 2003). REML makes use of all available data and retains cases with partial missing data. Repeated Level 1 measures were nested within students at Level 2. Intervention condition, basic demographics (e.g., gender, ethnicity, school site), measures of incoming achievement (e.g., prior GPA, standardized achievement test scores from the prior year), and total task time (for DPT analyses only) were variables at the student level. Individuals were permitted to deviate randomly from the mean intercept. For more details on the linear mixed effects model, refer to the online supplement.

Paralleling analyses in Studies 2 and 3, we first analyzed whether the intervention had a main effect, and then whether the intervention preferentially changed outcomes (e.g., beliefs, behaviors, and academic achievement) among lower-performing students. The Johnson-Neyman method was used to identify regions of significance.

Because the intervention's effects on beliefs and behaviors were not consistently moderated by prior achievement, and the intervention had no consistent effect on nonpractice-specific beliefs, these analyses are not reported in the main text. For full details on these analyses, as well as figures documenting the intervention's main effects on these outcomes, see the online supplement.

Data exclusions and missing data. In Study 4, as well as Study 5, many DPT sessions at Follow-up 1 were cancelled due to snow. Cancelled sessions were rescheduled for the first day following the snow day, which was often a half-day. Students who took part in rescheduled sessions experienced atypical testing conditions: On half-days, periods were often shorter than usual, and at several school sites, internet connectivity was either stalled or nonfunctional. In Study 4, the DPT session at Follow-up 1 was cancelled for approximately 40% of students. Because suboptimal experiences with the DPT at Follow-up 1 (e.g., the internet stalling, altered period schedules) aggravated students and colored their associations with the task at Follow-up 2, DPT data collected following snow cancellations were excluded from analyses. This same protocol was applied in Study 5: DPT data collected following snow cancellations were removed from analyses.

Approximately 7% of students did not complete self-report questionnaires at each of the three administrations. Participants with missing or excluded data (e.g., DPT data, self-reported beliefs) did not differ systematically from participants without missing data on any baseline variables (e.g., gender, ethnicity, second quarter GPA). For further details, see the online supplement.

Preliminary analyses. Randomization was successful. Treatment and control groups did not differ on baseline variables (for more details, see online supplement). Likewise, there was no between-condition difference in how interesting participants found the modules, $t(425) = 1.60$, $p = .11$, $d = -0.15$, 95% CI [−.34, .04].

Expectancy-value beliefs. At Follow-Up 1, students in the treatment condition reported stronger deliberate practice beliefs (practice-specific expectancy) than controls ($M_{\text{treatment}} = 75.33$, $M_{\text{control}} = 72.90$), $t(752) = 2.07$, $p = .04$, $d = .20$, 95% CI [.01, .40]. This effect persisted to Follow-Up 2, one month later ($M_{\text{treatment}} = 76.38$, $M_{\text{control}} = 73.37$), $t(752) = 2.32$, $p = .02$, $d = .25$, 95% CI [.05, .45]. Treatment participants also reported higher frustration tolerance (practice-specific cost, a component of value) than controls at Follow-up 1 ($M_{\text{treatment}} = 3.88$, $M_{\text{control}} = 3.54$), $t(750) = 5.93$, $p < .001$, $d = .68$, 95% CI [.47, .88], and Follow-Up 2 ($M_{\text{treatment}} = 3.77$, $M_{\text{control}} = 3.48$), $t(750) = 5.62$, $p < .001$, $d = .60$, 95% CI [.40, .80].

Deliberate practice behavior. At Follow-Up 1, 2–3 days following the intervention, students in the treatment condition did more deliberate practice than controls ($M_{\text{treatment}} = 944.80$, $M_{\text{control}} = 889.94$), $t(380) = 2.18$, $p = .03$, $d = .25$, 95% CI [.01, .52]. This effect persisted to Follow-Up 2, one month later ($M_{\text{treatment}} = 926.24$, $M_{\text{control}} = 861.50$), $t(380) = 2.68$, $p = .01$, $d = .29$, 95% CI [.01, .57].

Academic achievement. Students in the treatment condition earned higher fourth quarter GPAs than controls ($M_{\text{treatment}} = 0.09$, $M_{\text{control}} = -0.03$), $t(426) = 2.21$, $p = .03$, $d = .21$, 95% CI [.02, .40]. This effect was driven by lower-achievers, Condition × Prior Achievement interaction, $b = -1.14$, $SE = 0.47$, $t(415) = -2.43$, $p = .02$ (see Figure 4). A follow-up Johnson-Neyman analysis (Hayes, 2013; Johnson & Neyman, 1936) revealed that the effect of treatment was significant among lower-performing students (<0.12 *SD* above average on prior achievement).

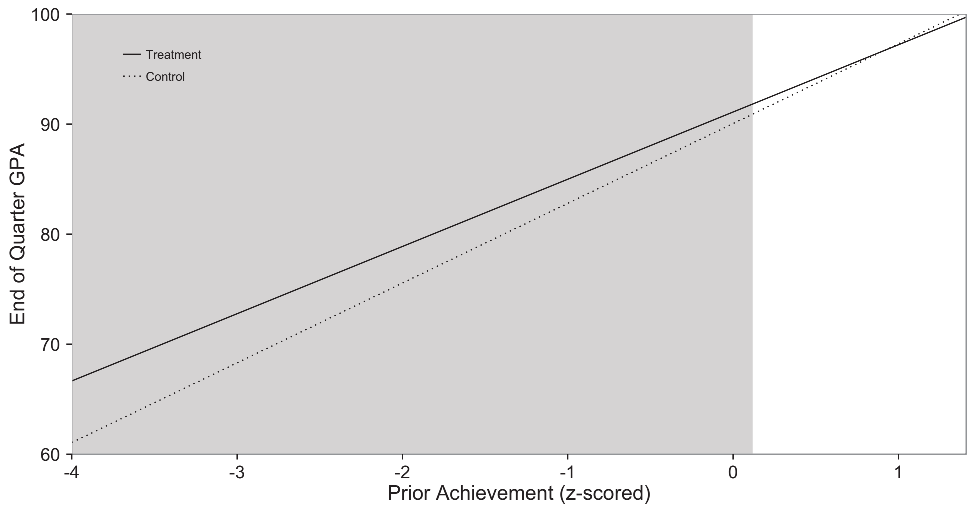

Figure 4. In Study 4, the intervention increased fourth quarter grades for all sixth graders, and especially among lower achievers. Shaded areas indicate regions of significance.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Because the intervention's effect on end-of-quarter achievement was driven by lower achievers, likely due to ceiling effects, we only expected changes in beliefs and behaviors to mediate the intervention's effect on achievement in this lower performing subgroup. Unfortunately, the current sample was underpowered to test for moderated mediation (Fritz & MacKinnon, 2007; for further details and exploratory analyses, see the online supplement).

Study 5

In Study 4, an expanded deliberate practice intervention delivered to sixth graders close to the start of the fourth quarter improved practice-specific expectancies and values, deliberate practice behavior, and fourth quarter achievement. In Study 5, we aimed to replicate these findings and test boundary conditions. Would the intervention have the same effect among seventh graders in a different district? Whereas in Study 4 we assessed beliefs and behaviors one month later, in Study 5 we conducted a follow-up assessment four months later. Testing the same intervention in a different district, with an older population, using an extended timeline, we probed the generalizability and longevity of the intervention's effects.

We introduced one final innovation in Study 5. Students were randomized to one of *three* conditions: full treatment, half treatment, or control. Full treatment participants received the same two modules as the treatment group in Study 4. Control participants received the same two modules as the control group in Study 4. The half treatment was a hybrid. These participants received the first treatment module (treatment condition, Part 1) followed by the second control module (control condition, Part 2). Thus, participants in the half treatment condition learned about deliberate practice, but did not receive the motivational content that targeted expectancies and values. We included this hybrid condition to parse the effects of the first treatment module from the second. Would teaching the tenets of deliberate practice suffice to change behavior (half treatment) or does deliberate practice need to be motivated (full treatment)?

Method

Participants. A total of 248 seventh-grade students at a public school district in the United States were invited to participate. Of these students, approximately 94% ($N = 232$) took the intervention. The remaining 6% of students either opted out, experienced technological malfunctions (e.g., computer crashed), or were absent. Administrators in Study 5 made every effort to schedule make-ups, thus minimizing the number of nonparticipants. Participants and nonparticipants did not differ on baseline variables (e.g., gender, ethnicity, first-quarter GPA). For details, see the online supplement.

Approximately 65.0% of participants were White, 13.8% Black, 10.3% Asian, and 10.3% Hispanic; 52.6% were female. Twelve percent of students qualified for free or reduced price meals.

Procedure.

Overview. Procedures were similar to those in Study 4. In the fall, students were administered a battery of self-report scales and the DPT during three periods over six school days. Approximately 8 weeks later, at the end of the second quarter, students completed the intervention during two periods spread across three school days. Because all classes in the district operated on a 2-day cycle (A-day/B-day), students participated every other day. One week after the intervention (Follow-up 1), and then again 4 months later (Follow-up 2), students retook the same battery of measures they completed preintervention. As in Study 4, students wore ear buds during all study sessions and privacy screens were attached to all computers.

For approximately 60% of students (B-day students at both middle schools, and A-day students at one middle school), the DPT was cancelled due to snow. Following the protocol described in Study 4 (refer there for details), DPT data collected during and after rescheduled sessions were excluded from analyses. The percentage of students with missing self-report data ranged from 4%–13% across the three data collections. Participants with missing data (e.g., DPT data, self-reported beliefs) did not differ from participants without missing data on any baseline variables (e.g.,

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

gender, ethnicity, first-quarter GPA). For further details, see the online supplement.

Intervention procedures and content. Students were randomized to one of three conditions: full treatment, half treatment, or control. Full treatment participants received the same two modules that the treatment group received in Study 4. Control participants took the same two modules that the control group took in Study 4. The half treatment received a mix: the first treatment module (treatment condition, Part 1, which taught the tenets of deliberate practice) followed by the second control module (control condition, Part 2, which discussed interests and achievement). Part 1 and Part 2 of the treatment and control modules are described in greater detail in Study 4.

Measures.

Self-report questionnaires. The self-report measures from Study 4 were readministered in Study 5. Once again, students took practice-specific measures of expectancy (deliberate practice beliefs at Follow-up 1, $\alpha = .57$; at Follow-up 2, $\alpha = .67$), and cost (frustration tolerance at preintervention, $\alpha = .50$; at Follow-up 1, $\alpha = .64$; at Follow-up 2, $\alpha = .57$), as well as nonpractice-specific measures of expectancy (growth mindset at preintervention, $\alpha = .67$; at Follow-up 1, $\alpha = .78$; at Follow-up 2, $\alpha = .85$; locus of control at preintervention, $\alpha = .45$; at Follow-up 1, $\alpha = .67$; at Follow-up 2, $\alpha = .65$), and cost (distress tolerance at preintervention, $\alpha = .85$; at Follow-up 1, $\alpha = .88$; at Follow-up 2, $\alpha = .89$). Due to time constraints, deliberate practice beliefs were not measured preintervention.

DPT, time focused, and total task time. Participants completed the same task described in Study 1.

GPA, standardized test scores, and basic demographics. Because the intervention was administered toward the end of the second quarter, we assessed the intervention's effect on third-quarter GPA, controlling for GPA from the prior quarter (first quarter). GPA from the first and third quarters, standardized achievement test scores from the prior year, as well as gender and ethnicity were obtained from school records.

Interest. As in Study 4, students rated how interesting they found each of the two modules, $r = .75$, $p < .001$, and these ratings were averaged to create a single measure of student interest.

Results and Discussion

Across reported analyses, the effect of half treatment did not differ systematically from the effect of full treatment or control. Results for half treatment participants are reported following results for full treatment participants.

Preliminary analyses. No baseline differences emerged between groups on any measured variables (for details, see online supplement). Likewise, a one-way analysis of variance revealed no between-condition differences in how interesting students found the modules, $F(2, 229) = 0.76$, $p = .47$, $\eta^2 = .01$.

Expectancy-value beliefs. At Follow-up 1, full treatment participants reported higher deliberate practice beliefs ($M_{\text{full treatment}} = 67.36$, $M_{\text{half treatment}} = 68.77$, $M_{\text{control}} = 62.55$), $t(208) = 2.19$, $p = .03$, $d = .36$, 95% CI [.03, .69], and frustration tolerance ($M_{\text{full treatment}} = 3.91$, $M_{\text{half treatment}} = 3.52$, $M_{\text{control}} = 3.51$), $t(396) = 2.05$, $p = .04$, $d = .32$, 95% CI [.01, .64] than controls. Neither the effect on deliberate practice beliefs ($M_{\text{full treatment}} = 75.73$, $M_{\text{half treatment}} = 76.95$, $M_{\text{control}} = 71.63$), $t(188) = 1.14$, $p = .26$, $d = .20$, 95% CI [−.14, .53] nor the effect on frustration tolerance, ($M_{\text{full treatment}} = 3.80$, $M_{\text{half treatment}} = 3.59$, $M_{\text{control}} = 3.64$), $t(396) = -0.11$, $p = .91$, $d = -.03$, 95% CI [−.36, .31] persisted to Follow-up 2.

Half treatment participants reported higher deliberate practice beliefs, $t(208) = 2.88$, $p = .004$, $d = .46$, 95% CI [.14, 79], but not higher frustration tolerance, $t(396) = -0.21$, $p = .83$, $d = -.02$, 95% CI [−.34, .29] at Follow-up 1. At Follow-up 2, half treatment participants did not differ from controls on deliberate practice beliefs, $t(188) = 1.48$, $p = .14$, $d = .25$, 95% CI [−.09, .59] or frustration tolerance, $t(396) = -0.64$, $p = .52$, $d = .11$, 95% CI [−.22, .45].

Deliberate practice behavior. At Follow-Up 1, students in the full treatment condition trended toward doing more deliberate practice than controls, but this effect was not statistically significant ($M_{\text{full treatment}} = 1312.40$, $M_{\text{half treatment}} = 1061.82$, $M_{\text{control}} = 1115.00$), $t(96) = 1.84$, $p = .07$, $d = .46$, 95% CI [−.19, 1.13]. At Follow-Up 2, 4 months later, there was no effect of treatment ($M_{\text{full treatment}} = 1232.10$, $M_{\text{half treatment}} = 1121.47$, $M_{\text{control}} = 1156.47$), $t(96) = 0.50$ $p = .61$, $d = .13$, 95% CI [−.54, .81].

For students in the half treatment condition, there were no significant main effects at Follow-up 1, $t(96) = -0.61$, $p = .55$, $d = -.13$, 95% CI [−.73, .47] or Follow-up 2, $t(96) = -0.45$, $p = .72$, $d = -.08$, 95% CI [−.73, .56].

Academic achievement. There was no main effect of full treatment ($M_{\text{full treatment}} = 0.08$, $M_{\text{half treatment}} = -0.02$, $M_{\text{control}} = -0.05$), $t(217) = 1.47$, $p = .14$, $d = .25$, 95% CI [−.03, .52], nor was there a Full Treatment × Prior Achievement interaction, $b = -0.15$, $SE = 0.09$, $t(217) = -1.63$, $p = .11$. Nevertheless, because the interaction was in the same direction as the interaction in Study 4, we probed it to see if the same pattern of results emerged. It did. A follow-up Johnson-Neyman analysis (Hayes, 2013; Johnson & Neyman, 1936) revealed that the effect of full treatment was significant among lower performing students (<-0.31 *SD* below average on prior achievement; see Figure 5).

For students in the half treatment condition, there was no significant main effect of intervention, $t(231) = 0.27$, $p = .79$, $d = .06$, 95% CI [−.21, .33]. Nor was the intervention effect moderated by prior achievement: Half Treatment × Prior Achievement interaction, $b = -0.16$, $SE = 0.11$, $t(231) = -1.48$, $p = .14$.

Because the intervention's effect on end-of-quarter achievement was driven by lower achievers, we only expected changes in beliefs and behaviors to mediate the intervention's effect in this lower performing subgroup. Yet, as in Study 4, the sample was underpowered to test for moderated mediation (Fritz & MacKinnon, 2007; for exploratory analyses, see the online supplement).

Synthesis of Study Results

Because the size and significance of intervention effects on achievement varied across the four intervention studies—two of the four main effects, and three of the four Condition × Prior Achievement interactions reached significance (the fourth interaction approached significance)—we conducted two meta-analyses using the "metafor" package in R. In the first meta-analytic model, we tested whether there was a significant mean difference on postintervention achievement between the treat-

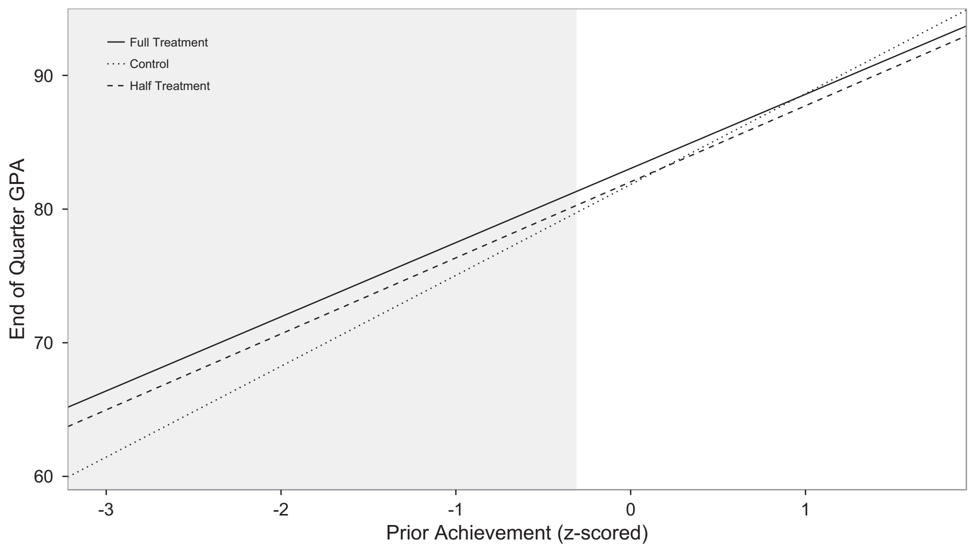

Figure 5. In Study 5, the full treatment condition increased third-quarter grades for lower achieving seventh graders. Shaded areas indicate regions where full treatment differed from the control condition (half treatment did not differ from control at any values of the moderator).

ment and control groups across the four interventions. Effect sizes were calculated using Hedges' *g*. A mean effect size (g_+) was calculated by weighting each study effect size by its inverse variance and averaging across the weighted estimates (Lipsey & Wilson, 2001). Given the small number of studies in the meta-analysis, we combined means using a fixed-effects model. We found a small but significant mean effect ($g_+ = 0.20$; $p = .002$, 95% CI [.08, .33]) and found little evidence of heterogeneity across studies, $Q(3) = 1.02$, $p = .80$. This suggests that, on average, the intervention improved achievement outcomes (see Figure 6a). We then ran a random-effects model as a sensitivity analysis using the DerSimonian-Laird estimator. The random-effects model showed little evidence of between-study variability ($I^2 = 0\%$) and yielded identical mean effect size estimates and confidence limits to the hundreds decimal place.

Next, we ran a meta-analytic model testing whether the mean interaction coefficient for the Condition × Prior Achievement interaction was significant across the four studies. Regression coefficients for the interaction effects were weighted by their inverse variance and pooled using a fixed-effects model. The mean coefficient for the Condition × Prior Achievement interaction was significant ($b_+ = -.18$; $p < .001$, 95% CI [$-.27$, $-.10$]) and we found little hetereogeneity across studies, $Q(3) = 4.25$, $p = .24$. This suggests that the magnitude of the

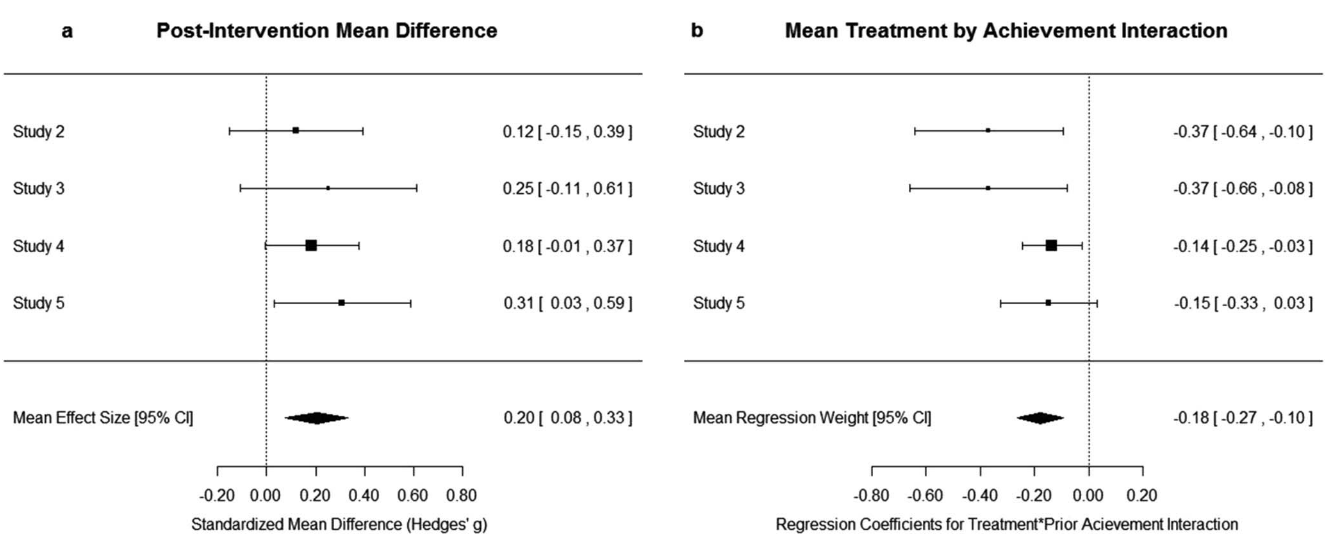

Figure 6. On average, across Studies 2–5, mean achievement scores were 0.20 standard deviations higher in the treatment conditions than control conditions at postintervention (Panel a). Additionally, the weighted mean coefficient for the Treatment × Prior Achievement interaction was significantly different than 0 (Panel b).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

intervention effect differed as a function of prior achievement levels (see Figure 6b). A random-effects model using the DerSimonian-Laird estimator yielded highly similar results ($b_+ = -.20$; $p = .0004$, 95% CI [−.31, −.09]).

General Discussion

In the current investigation, a psychologically "wise" deliberate practice intervention improved expectancy-value beliefs, deliberate practice, and academic achievement among nonexperts. In Study 1, we validated a behavioral measure of deliberate practice and confirmed the importance of deliberate practice to nonexpert achievement: Middle school students who did more deliberate practice earned higher grades. Next, across four longitudinal, randomized-controlled, field experiments, we intervened. In Study 2, a brief intervention improved the rate at which lower-achieving fifth and sixth graders mastered math in Khan Academy. In Study 3, an adapted version of the same intervention delivered to undergraduates raised end-of-semester grades, particularly for lower performing students. In Study 4, an expanded intervention improved sixth graders' practice-specific expectancy-value beliefs, deliberate practice behavior, and academic achievement, with the strongest effects on achievement found among lower performing students. A similar pattern of results emerged among seventh graders in Study 5. In sum, across four intervention studies we analyzed effects on three key outcomes: beliefs, deliberate practice, and achievement. The intervention changed beliefs and behaviors for 1 month (Study 4) but not 4 (Study 5), and improved achievement—particularly among lower achievers—over one academic quarter.

Positive effects occurred regardless of whether the intervention was delivered to students in middle school (Studies 2, 4, 5) or college (Study 3). Likewise, the intervention affected a range of achievement-related outcomes: among lower performing students, the intervention improved math performance in Khan Academy (Study 2), end-of-course grades (Study 3), end-of-semester GPA (Study 3), and end-of-quarter GPA (Study 4; the effect was in the same direction but not significant in Study 5). To synthesize effects on achievement, we conducted a meta-analytic estimate of the main effect and the interaction effect across all four studies. We found a small but significant main effect and a significant Condition × Prior Achievement interaction.

The treatment condition proved beneficial above and beyond two control conditions, one that instructed students in standard study skills (Studies 2 and 3), and the other that discussed deep interests, generalized motivation, and high achievement (Studies 4 and 5). Across studies, teachers were blind to hypotheses, condition assignment, and intervention content. Thus, we believe we can rule out the possibility that gains in achievement were confounded by teacher expectancy.

Contributions of the Present Research

On a theoretical level, this research supports the heretofore untested hypothesis that deliberate practice submits to the same motivational principles that govern ordinary effortful behavior. Across a series of randomized-controlled experiments, wise interventions targeting expectancies and values motivated deliberate practice and improved achievement among nonexpert students. Past research documents that expectancy and value are associated with persistence and performance (Eccles, 1983; Eccles et al., 1984; Eccles & Wigfield, 2002; Wigfield & Eccles, 2000); yet to our knowledge, the current investigation is the first to demonstrate that manipulating these variables can increase deliberate practice. Though our findings suggests that exceptional levels of grit and passion are not required to motivate deliberate practice, it appears some form of motivation is. In Study 5, students who were taught the tenets of deliberate practice without an accompanying motivational lesson (targeting self-relevant expectancies and values) did not improve their deliberate practice behavior or their achievement.

The DPT represents an important methodological advance. This novel task measures deliberate practice among novice students. Whereas standard self-report measures of deliberate practice (e.g., retrospective estimates, diaries) rely on participants to accurately report the quality of their practice—requiring metacognitive insight that is present among experts but not necessarily among nonexperts—the DPT circumvents this need by assessing behavior objectively. Task measures have their own limitations (e.g., practice effects; Duckworth, Yeager, & Bryk, 2014). However, we believe the DPT provides a useful counterpart to self-report measures. In the present investigation the DPT featured math problems, but other content could easily be substituted to adapt the task to measure deliberate practice in other domains (e.g., science, literacy).

Finally, the present investigation makes a timely practical contribution. A brief, wise intervention improved the academic performance of students in middle school and college. Whereas the intervention raised achievement for all students in some studies (Studies 3 and 4), the intervention consistently improved the academic performance of lower achievers (Studies 2, 3, 4, and 5).

Limitations and Future Directions

Limitations of the current investigation suggest several profitable directions for future research. First and foremost, our intervention did not turn nonexperts into experts. World-class experts average several hours of deliberate practice per day, over the course of years (Ericsson, 2006). Contrast this with students in the current investigation who increased their deliberate practice for 1 month (Study 4) but not 4 (Study 5). As opposed to transforming novices into experts, what these studies provide is proof of concept: Brief, motivational interventions can promote short-term increases in the amount of deliberate practice nonexperts voluntarily perform. Future research is needed to determine precisely how long effects last, the mechanism(s) by which they attenuate, and whether booster sessions can halt this decline.

Future research is also needed to determine whether the intervention can raise achievement for *all* students. In the current investigation, the intervention preferentially helped lower achievers (Studies 2, 3, 4, and 5). We believe this was due to ceiling effects. Measures of achievement could not capture improvement among students who were excelling prior to the intervention. Because almost all individuals in nonexpert settings do less deliberate practice than they are capable of, future studies ought to assess achievement outcomes that do not have restriction on range. On such outcomes, we expect the intervention would benefit all students.

How can future interventions be strengthened? One possibility is to redesign the intervention to target the individual in his broader con-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

text. The present intervention taught students to execute deliberate practice absent teacher involvement—participation involved interacting with a computer. In the real world, even experts execute deliberate practice with the support of experienced others (e.g., coaches). It will be interesting for future studies to measure the effects of more expansive interventions—those which address not just students but also teachers, the grading system, and the broader school culture.

A second possibility is to see whether combining the present intervention with other active ingredients has synergistic effects. Despite obvious similarities between the present intervention and growth mindset interventions—both target achievement-related outcomes and aim to recalibrate attitudes to challenge—the present intervention had no reliable effect on self-reported growth mindset. Given this nonoverlap, it is possible that the present intervention, combined with a growth mindset intervention, would have an additive, even multiplicative, effect.

The various parts of the present, multifaceted intervention still need to be systematically disentangled. Deliberate practice encompasses a suite of behaviors (e.g., setting specific goals, seeking out feedback, working on weaknesses). Instructing students on any of these behaviors may have driven the results. Feedback, for example, is robustly associated with learning and achievement (Bangert-Drowns et al., 1991; Hattie & Timperley, 2007). Likewise, self-testing, an activity consistent with deliberate practice, independently promotes learning and performance (Roediger & Karpicke, 2006). Against these specific possibilities, up to one month following the intervention, students were more likely to choose to challenge themselves while working (as measured by the DPT), a behavior that does not directly follow from learning to self-test or seek feedback. Nevertheless, more investigations like Study 5 that parse the intervention to identify active ingredients are warranted.

For a number of reasons, we also recommend replications. One important reason to conduct replications is because the present intervention was, to some degree, customized. Focus groups were drawn from the schools in which the interventions were run. It would be interesting to know whether an intervention that has *not* been customized to the participating student body leads to similar results. In addition, the present studies suffered from a number of irregularities (e.g., missing data, snow day cancellations). Replications, particularly in samples that are more diverse, would provide further evidence of generalizability.

Finally, we recommend replications in larger samples, which would permit tests of moderated mediation. Because the intervention only consistently improved achievement for lower performing students, one would expect to find the intervention's effect on achievement mediated by a change in beliefs and behaviors in this subsample. Moderated mediation analyses, which the current studies were underpowered to perform, would permit a more conclusive determination of whether the beliefs and behaviors that changed postintervention explain the achievement gains of lower-performing students. A current frontier in the study of wise interventions is to identify the precise pathways through which they generate enduring effects (Walton, 2014; Yeager & Walton, 2011). Thus, mechanisms that account for the current intervention, like those of other wise interventions, require continued exploration.

Conclusion

Deliberate practice was first investigated as a predictor of excellence among experts (Ericsson et al., 1993). To a large degree, the domain of expertise is where the study of deliberate practice has remained. We hope our findings goad researchers and educators alike to wonder, why? Why *not* teach this method of continuous improvement to, well, everyone? In the present investigation, a brief intervention anchored in psychological theory changed beliefs, deliberate practice, and academic achievement for up to one academic quarter. Future research is needed to determine whether these effects can be strengthened, and if so, to what extent. Can deliberate practice interventions transform novices into experts, or merely make them expert-like? Answers to these questions will reveal the full extent to which deliberate practice, and excellence itself, can be encouraged among nonexperts.

References

Aronson, E. (1999). The power of self-persuasion. *American Psychologist, 54,* 875–884. http://dx.doi.org/10.1037/h0088188

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38,* 113–125. http://dx.doi.org/10.1006/jesp.2001.1491

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64,* 359–372. http://dx.doi.org/10.1037/h0043445

Autin, F., & Croizet, J.-C. (2012). Improving working memory efficiency by reframing metacognitive interpretation of task difficulty. *Journal of Experimental Psychology: General, 141,* 610–618. http://dx.doi.org/10.1037/a0027478

Bangert-Drowns, R. L., Kulik, C. L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61,* 213–238. http://dx.doi.org/10.3102/00346543061002213

Battle, E. S. (1965). Motivational determinants of academic task persistence. *Journal of Personality and Social Psychology, 2,* 209–218. http://dx.doi.org/10.1037/h0022442

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,* 246–263. http://dx.doi.org/10.1111/j.1467-8624.2007.00995.x

Bonneville-Roussy, A., Lavigne, G. L., & Vallerand, R. J. (2011). When passion leads to excellence: The case of musicians. *Psychology of Music, 39,* 123–138. http://dx.doi.org/10.1177/0305735609352441

Bryan, W. L., & Harter, N. (1897). Studies in the physiology and psychology of the telegraphic language. *Psychological Review, 4,* 27–53. http://dx.doi.org/10.1037/h0073806

Cialdini, R. B. (2003). Crafting normative messages to protect the environment. *Current Directions in Psychological Science, 12,* 105–109. http://dx.doi.org/10.1111/1467-8721.01242

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology, 58,* 1015–1026. http://dx.doi.org/10.1037/0022-3514.58.6.1015

Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science, 324,* 400–403. http://dx.doi.org/10.1126/science.1170769

Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology, 112,* 558–577. http://dx.doi.org/10.1037/0021-843X.112.4.558

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Crandall, V. C. (1969). Sex differences in expectancy of intellectual and academic reinforcement. In C. P. Smith (Ed.), *Achievement-related motives in children* (pp. 11–45). New York, NY: Russell Sage Foundation.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science, 332,* 862–864. http://dx.doi.org/10.1126/science.1201783

Duckworth, A. L., Kirby, T., Tsukayama, E., Berstein, H., & Ericsson, K. A. (2011). Deliberate practice spells success: Why grittier competitors triumph at the National Spelling Bee. *Social Psychological & Personality Science, 2,* 174–181. http://dx.doi.org/10.1177/1948550610385872

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92,* 1087–1101. http://dx.doi.org/10.1037/0022-3514.92.6.1087

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short grit scale (grit-s). *Journal of Personality Assessment, 91,* 166–174. http://dx.doi.org/10.1080/00223890802634290

Duckworth, A. L., Yeager, D. S., & Bryk, A. S. (2014). *Measurement matters: Assessing attributes other than cognitive ability*. Manuscript in preparation.

Dweck, C. S., Chiu, C.-Y., & Hong, Y.-Y. (1995). Implicit theories and their role in judgments and reactions: A world from two perspectives. *Psychological Inquiry, 6,* 267–285. http://dx.doi.org/10.1207/s15327965pli0604_1

Eccles, J. S. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and socialogical approaches* (pp. 283–331). San Francisco, CA: Freeman and Company.

Eccles, J. S., Midgley, C., & Adler, T. (1984). Grade-related changes in the school environment: Effects on achievement motivation. *The Development of Achievement Motivation, 3,* 283–331.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53,* 109–132. http://dx.doi.org/10.1146/annurev.psych.53.100901.135153

Eccles, J., 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,* 830–847. http://dx.doi.org/10.2307/1131221

Eisenberger, R. (1992). Learned industriousness. *Psychological Review, 99,* 248–267. http://dx.doi.org/10.1037/0033-295X.99.2.248

Ericsson, K. A. (1996). The acquisition of expert performance: An introduction to some of the issues. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 1–50). Mahwah, NJ: Erlbaum.

Ericsson, K. A. (2002). Attaining excellence through deliberate practice: Insights from the study of expert performance. In M. Ferrari (Ed.), *The pursuit of excellence through education*. The educational psychology series (pp. 21–55). Mahwah, NJ: Erlbaum Publishers. http://dx.doi.org/10.1002/9780470690048.ch1

Ericsson, K. A. (2003). How the expert performance approach differs from traditional approaches to expertise in sport. *Expert performance in sports—Advances in research on sport expertise* (pp. 371–402). Champaign, IL: Human Kinetics Publishers.

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 685–706). Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511816796.038

Ericsson, K. A. (2007). Deliberate practice and the modifiability of body and mind: Toward a science of the structure and acquisition of expert and elite performance. *International Journal of Sport Psychology, 38,* 4–34.

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine, 15,* 988–994. http://dx.doi.org/10.1111/j.1553-2712.2008.00227.x

Ericsson, K. A. (2009). Enhancing the development of professional performance: Implications from the study of deliberate practice. In K. A. Ericsson (Ed.), *The development of professional expertise: Toward measurement of expert performance and design of optimal learning environments* (pp. 405–431). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511609817.022

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49,* 725–747. http://dx.doi.org/10.1037/0003-066X.49.8.725

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363–406. http://dx.doi.org/10.1037/0033-295X.100.3.363

Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2007). The making of an expert. *Harvard Business Review, 85,* 114–121, 193.

Ericsson, K. A., & Ward, P. (2007). Capturing the naturally occurring superior performance of experts in the laboratory: Toward a science of expert and exceptional performance. *Current Directions in Psychological Science, 16,* 346–350. http://dx.doi.org/10.1111/j.1467-8721.2007.00533.x

Feather, N. T. (1982). Expectancy-value approaches: Present status and future directions. In N. T. Feather (Ed.), *Expectations and actions: Expectancy-value models in psychology* (pp. 395–420). Hillsdale, NJ: Lawrence Erlbaum.

Feingold, A. (2015). Confidence interval estimation for standardized effect sizes in multilevel and latent growth modeling. *Journal of Consulting and Clinical Psychology, 83,* 157–168. http://dx.doi.org/10.1037/a0037721

Fritz, M. S., & Mackinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science, 18,* 233–239. http://dx.doi.org/10.1111/j.1467-9280.2007.01882.x

Garcia, J., & Cohen, G. L. (2012). A social psychological approach to educational intervention. In E. Shafir (Ed.), *Behavioral foundations of public policy* (pp. 329–347). Princeton, NJ: Princeton University Press.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research, 35,* 472–482. http://dx.doi.org/10.1086/586910

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology, 24,* 645–662. http://dx.doi.org/10.1016/j.appdev.2003.09.002

Graham, M. (1998). I am a dancer. In A. Carter (Ed.), *The Routledge dance studies reader* (pp. 66–71). New York, NY: Routledge.

Gray, W. D. (2000). The nature and processing of errors in interactive behavior. *Cognitive Science, 24,* 205–248. http://dx.doi.org/10.1207/s15516709cog2402_2

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

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press.

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326,* 1410–1412. http://dx.doi.org/10.1126/science.1177067

Jamieson, J. P., Mendes, W. B., Blackstock, E., & Schmader, T. (2010). Turning the knots in your stomach into bows: Reappraising arousal improves performance on the GRE. *Journal of Experimental Social Psychology, 46,* 208–212. http://dx.doi.org/10.1016/j.jesp.2009.08.015

John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). New York, NY: Guilford Press.

Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. In J. Neyman & E. S. Pearson (Eds.), *Statistical research memoirs* (pp. 57–93). New York, NY: Taylor & Francis.

Kessler, D. O., Auerbach, M., Pusic, M., Tunik, M. G., & Foltin, J. C. (2011). A randomized trial of simulation-based deliberate practice for infant lumbar puncture skills. *Simulation in Healthcare, 6,* 197–203. http://dx.doi.org/10.1097/SIH.0b013e318216bfc1

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General, 139,* 665–682. http://dx.doi.org/10.1037/a0020198

Lewin, K. (1952). Group decision and social change. In G. E. Swanson, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in social psychology* (2nd ed., pp. 330–344). New York, NY: Holt.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (Vol. 49). Thousand Oaks, CA: Sage.

Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science, 25,* 1608–1618. http://dx.doi.org/10.1177/0956797614535810

McGaghie, W. C., Issenberg, S. B., Cohen, M. E. R., Barsuk, J. H., & Wayne, D. B. (2011). Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. *Academic Medicine: Journal of the Association of American Medical Colleges, 86,* 706.

Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology, 82,* 60–70. http://dx.doi.org/10.1037/0022-0663.82.1.60

Nagengast, B., Trautwein, U., Kelava, A., & Lüdtke, O. (2013). Synergistic effects of expectancy and value on homework engagement: The case for a within-person perspective. *Multivariate Behavioral Research, 48,* 428–460. http://dx.doi.org/10.1080/00273171.2013.775060

Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science, 26,* 784–793. http://dx.doi.org/10.1177/0956797615571017

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95,* 667–686. http://dx.doi.org/10.1037/0022-0663.95.4.667

Plant, E. A., Ericsson, K. A., Hill, L., & Asberg, K. (2005). Why study time does not predict grade point average across college students: Implications of deliberate practice for academic performance. *Contemporary Educational Psychology, 30,* 96–116. http://dx.doi.org/10.1016/j.cedpsych.2004.06.001

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135,* 322–338. http://dx.doi.org/10.1037/a0014996

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42,* 185–227. http://dx.doi.org/10.1080/00273170701341316

Roediger, H. L., III, & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1,* 181–210. http://dx.doi.org/10.1111/j.1745-6916.2006.00012.x

Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. New York, NY: McGraw-Hill.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195152968.001.0001

Smith, V. L., & Walker, J. M. (1993). Monetary rewards and decision cost in experimental economics. *Economic Inquiry, 31,* 245–261. http://dx.doi.org/10.1111/j.1465-7295.1993.tb00881.x

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72,* 271–324. http://dx.doi.org/10.1111/j.0022-3506.2004.00263.x

Tsay, C. J., & Banaji, M. R. (2011). Naturals and strivers: Preferences and beliefs about sources of achievement. *Journal of Experimental Social Psychology, 47,* 460–465. http://dx.doi.org/10.1016/j.jesp.2010.12.010

Vallerand, R. J., Mageau, G. A., Elliot, A. J., Dumais, A., Demers, M-A., & Rousseau, F. (2008). Passion and performance attainment in sport. *Psychology of Sport and Exercise, 9,* 373–392. http://dx.doi.org/10.1016/j.psychsport.2007.05.003

Vallerand, R. J., Salvy, S.-J., Mageau, G. A., Elliot, A. J., Denis, P. L., Grouzet, F. M. E., & Blanchard, C. (2007). On the role of passion in performance. *Journal of Personality, 75,* 505–533. http://dx.doi.org/10.1111/j.1467-6494.2007.00447.x

Walton, G. M. (2014). The new science of wise psychological interventions. *Current Directions in Psychological Science, 23,* 73–82. http://dx.doi.org/10.1177/0963721413512856

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science, 331,* 1447–1451. http://dx.doi.org/10.1126/science.1198364

Wellborn, J. G., Connell, J. P., & Skinner, E. A. (1989). *The students perceptions of control questionnaire (SPOCQ): Academic domain*. Rochester, NY: University of Rochester.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25,* 68–81. http://dx.doi.org/10.1006/ceps.1999.1015

Wilson, T. D. (2011). *Redirect: The surprising new science of psychological change*. New York, NY: Little, Brown and Company.

Wilson, T. D., & Linville, P. W. (1982). Improving the academic performance of college freshmen: Attribution therapy revisited. *Journal of Personality and Social Psychology, 42,* 367–376. http://dx.doi.org/10.1037/0022-3514.42.2.367

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. http://dx.doi.org/10.1080/00461520.2012.722805

Yeager, D. S., Henderson, M. D., Paunesku, D., Walton, G. M., D'Mello, S., Spitzer, B. J., & Duckworth, A. L. (2014a). Boring but important: A self-transcendent purpose for learning fosters academic self-regulation. *Journal of Personality and Social Psychology, 107,* 559–580. http://dx.doi.org/10.1037/a0037637

Yeager, D. S., Johnson, R., Spitzer, B. J., Trzesniewski, K. H., Powers, J., & Dweck, C. S. (2014b). The far-reaching effects of believing people can change: Implicit theories of personality shape stress, health, and achievement during adolescence. *Journal of Personality and Social Psychology, 106,* 867–884. http://dx.doi.org/10.1037/a0036335

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81,* 267–301. http://dx.doi.org/10.3102/0034654311405999

Received March 22, 2015
Revision received August 11, 2015
Accepted September 5, 2015 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.