



HHS Public Access

Author manuscript

J Pers Soc Psychol. Author manuscript; available in PMC 2022 January 01.

Published in final edited form as:

J Pers Soc Psychol. 2021 January ; 120(1): 1–15. doi:10.1037/pspa0000246.

Psychological Affordances Help Explain Where a Self-Transcendent Purpose Intervention Improves Performance

Stephanie L. Reeves,

Department of Psychological and Brain Sciences, Indiana University

Marlone D. Henderson,

Department of Psychology, The University of Texas at Austin

Geoffrey L. Cohen,

Department of Psychology and Graduate School of Education, Stanford University

Rebecca R. Steingut,

Independent Practice, New York, New York

Quinn Hirschi,

Department of Psychology, University of Virginia

David S. Yeager

Department of Psychology, The University of Texas at Austin.

Abstract

Lay theory interventions instill situation-general ways of thinking, often using short reading and writing exercises, and they have led to lasting changes in behavior and performance in a wide variety of policy domains. Do they work in all contexts? We suggest that lay theory intervention effects depend on *psychological affordances*, which are defined as cues that allow individuals to view a lay theory as legitimate and adaptive in that context. The present research directly and experimentally tested this hypothesis using the example of a “purpose for learning” lay theory intervention, which taught the lay theory that school is a place to develop skills that allow one to make progress toward self-transcendent aims. A double-blind 2 (student purpose intervention) \times 2 (purpose-affording note) field experiment was conducted in a relatively low-performing public middle school in the United States. Students first received a web-based purpose for learning lay theory intervention (or a control activity), and 2 weeks later attended a class in which an assignment was accompanied by a purpose-affording note that was hand-written by a teacher (or a control note). Results showed that the purpose lay theory intervention increased performance on an English class writing assignment, but only when it was accompanied by a purpose-affording note. Exploratory analyses revealed that the effects of the manipulations were apparent among students who were at greater risk for poor performance in the class: nonnative English-speaking

Correspondence concerning this article should be addressed to Stephanie L. Reeves, Department of Psychological and Brain Sciences, Indiana University, 1101 East 10th Street, Bloomington, IN 47405-7007, or to David S. Yeager, Department of Psychology, The University of Texas at Austin, 108 East Dean Keeton Street, Austin, TX 78712. stepreev@iu.edu or dyeager@utexas.edu.

Supplemental materials: <http://dx.doi.org/10.1037/pspa0000246.supp>

students. Thus short, online lay theory interventions may reduce performance gaps, provided that the contexts afford the opportunity for the proffered lay theory to seem legitimate and adaptive.

Keywords

motivation; purpose; wise interventions; lay theories; behavior change

One of the most striking developments in social psychology in recent years has been the finding that short, online interventions can cause lasting improvements in people's behavior and welfare across a wide variety of important policy domains (see Walton & Wilson, 2018). These interventions do not change the objective situation, but instead persuade people to change their "lay theories," which are defined as loosely articulated but coherent worldviews that guide people's construals of ambiguous situations and, thereby, shape motivation, and behavior (Gelman, 2003; Gopnik & Meltzoff, 1997; Molden & Dweck, 2006). Recently, large-scale field experiments (e.g., Yeager et al., 2019) have shown that lay theories and subsequent behavior can be changed through brief exercises that use persuasive techniques that are well-known to social psychologists (i.e., strong arguments, Petty & Cacioppo, 1986; descriptive norms, Miller & Prentice, 2016; and self-persuasion, Aronson, 1999). In the academic domain, examples of lay theory interventions (also called "mindset" or "wise" interventions) include growth mindset (Blackwell, Trzesniewski, & Dweck, 2007; Paunesku et al., 2015; Yeager et al., 2019), sense of belonging (Walton & Cohen, 2007, 2011; Yeager, Walton, et al., 2016), relevance (Harackiewicz, Canning, Tibbetts, Priniski, & Hyde, 2016; Hulleman, Godes, Hendricks, & Harackiewicz, 2010; Hulleman & Harackiewicz, 2009), wise feedback (Yeager, Purdie-Vaughns, et al., 2014), and more (see Destin & Oyserman, 2010).

Lay theory interventions have attracted attention in part because they can be brief (<1 hr) and delivered online, which means that they can be delivered widely and, therefore, might address persistent policy problems. This potential for wide-scale use in entire populations (e.g., whole schools, states, or nations); however, highlights a need to understand where lay theory interventions do and do not work. No intervention, no matter how well-designed, will have the same effects for all people in all contexts. Critically, intervention effect heterogeneity is not a limitation; it is an opportunity to learn about mechanism. If we can pinpoint *where and under what conditions* a lay theory has its effects on behavior, we can have a better understanding of why it works.

Here we report a field experiment designed to understand how a lay theory intervention depends on context. We randomly assigned adolescents to a lay theory intervention or a control, and we crossed this with a manipulation of a contextual factor that should afford the adolescent's lay theory in the situation, to advance theories of the intervention's mechanisms.

The Self-Transcendent Purpose for Learning Intervention: Why Does It Work?

The lay theory intervention that we examined here was the *self-transcendent purpose for learning* lay theory (Yeager, Henderson, et al., 2014), hereafter called the “purpose” lay theory for short. The purpose lay theory involves the presumption that teachers give students homework and other assignments in school because they want students to learn and build skills that they can use to have a positive impact on something beyond-the-self (family, community, and the world) and also to obtain a benefit to the self (e.g., to have an interesting and fulfilling career one day; Yeager, Henderson, et al., 2014). Purpose is typically contrasted to a more instrumental lay theory of school (i.e., that schooling is only about achieving high-enough grades or acquiring credentials), or to a simple focus on learning and skill-development without connecting these goals to students’ broader prosocial goals. In previous randomized experiments, an intervention that instilled the purpose lay theory increased learners’ self-regulation (in a laboratory study) and grades (in a field study) relative to control group students (Yeager, Henderson, et al., 2014).

How could a short purpose intervention have effects that persist over time? This has been puzzling (see Miller, Dannals, & Zlatev, 2017), in part because participants presumably do not keep the theory-induction message vividly in working memory for hours or days, let alone weeks or months (Walton & Cohen, 2011). *Recursive process* models have been proposed as a solution to this puzzle (Cohen, Garcia, Apfel, & Master, 2006; Walton & Wilson, 2018; Yeager & Walton, 2011). Recursive process models emphasize that, after receiving a lay theory intervention, individuals can come to construe specific situations, such as a difficult or boring homework assignment, in light of their new and more adaptive lay theory. By applying their situation-general theory to a specific situation, they experience greater motivation and engagement. The positive changes that result from greater motivation and engagement (e.g., higher grades, more classroom participation, positive reinforcement from teachers, etc.) go on to reinforce the lay theory. Although studies have demonstrated that lay theory interventions *can* initiate recursive processes under some conditions (Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009; Walton & Cohen, 2011), few have examined when and where a relatively short intervention continues to affect motivation and behavior over time.

Lay Theory Interventions Depend on Psychological Affordances

A key assumption of recursive process models—and the starting point for the present research—is that people actually encounter situational cues that encourage them to put their lay theories into practice. This assumption is critical because lay theory interventions give people “working hypotheses” (Walton & Wilson, 2018) about how to interpret ambiguous situations, but people still need to decide whether that working hypothesis applies *here, in this place*. In particular, people may rely on the *psychological affordances* of the context when deciding whether to apply lay theories in specific situations (see Gibson, 1977; Walton & Yeager, 2020).

Affordances are features of a situation that invite or “pull for” a particular interpretation and response. Psychologists often think of affordances in terms of physical objects (Gibson, 1977). For instance, a handle on a round, hollowed object invites the perception that “this can be drunk from” and a corresponding response if one is thirsty, but a person seeing the same object without a handle may not think about taking a sip. Similarly, psychological affordances refer to those features of a social environment that make a person’s prior lay theory accessible, legitimate, and adaptive. Psychological affordances, therefore, explain why the effects of social psychological interventions can be powerful under some circumstances but always depend on the context (Cohen, Garcia, & Goyer, 2017; Cohen & Sherman, 2014).

To illustrate our model, imagine a parent who has successfully instilled the purpose lay theory in a child. If the student enters a classroom context in which teachers frequently explain the relevance of the coursework and invite students to see the connections between stronger skills and a meaningful life, then the purpose lay theory will be accessible and useful in that particular classroom and the child will have evidence that the parent’s messages were legitimate. However, suppose the same adolescent had instead attended a school in which teachers never explained ways to make the schoolwork relevant to personal goals, and the work could not plausibly be viewed as preparation for a meaningful contribution. Presumably, students are not blind to ongoing social realities that fail to align with their lay theories. Despite the support for a purpose the student received at home, the student may not act on the purpose lay theory when teachers—who are the authority in the classroom—structure the norms and assignments in a way that does not afford the chance to act on the purpose lay theory.

The present research tests the moderating effect of psychological affordances directly and experimentally, for the first time, using a 2 (student lay theory intervention) \times 2 (affordance of a lay theory) double-blind, randomized field experiment. Our study examined whether the efficacy of the purpose for learning intervention (see Paunesku et al., 2015; Yeager, Henderson, et al., 2014) was moderated by the *purposeful affordances* in the academic context, communicated to students by their teacher. Purposeful affordances are defined here as cues that support students’ perception that classroom assignments are opportunities to build skills that will help them to pursue and develop their purposes for learning. Thus, purposeful affordances are cues that signal to students that the purpose lay theory is legitimate and adaptive in a particular classroom context. Testing the model of psychological affordances in this manner has implications not only for theories of how best to motivate behavior change, but also for recent discussions about the importance of context in predicting whether and when the effects of a lay theory intervention replicate (Yeager et al., 2019).

Previous Tests of Psychological Affordance Interactions

Until recently, research had not tested the assumption that the effects of lay theory interventions in education vary systematically as a function of psychological affordances. Two new studies have begun to provide empirical evidence at a large scale, and these have set the stage for the present investigation (for a review, see Walton & Yeager, 2020).

First among these is the *National Study of Learning Mindsets* (NSLM; Yeager et al., 2019). The NSLM is a longitudinal randomized controlled trial evaluating a growth mindset of intelligence intervention—that teaches the lay theory that intelligence is malleable—with >12,000 students in a representative sample of 65 U.S. high schools. Analyses discovered a Lay Theory × Psychological Affordances interaction, such that the intervention effects among lower-achieving students were larger when the context provided more opportunities to put the growth mindset into practice (i.e., when peers showed a positive norm of challenge-seeking). Yeager et al. (2019) interpreted the results as evidence that unsupportive school norms discouraged students from putting the growth mindset lay theory into practice, undermining its lasting effects.

The second finding comes from the *College Transition Collaborative* (CTC), an unpublished randomized, controlled trial testing the effects of a social belonging lay theory intervention on college persistence among $N > 39,000$ incoming freshmen from 22 colleges and universities (Walton, 2019). This intervention encouraged students to view difficulties in the transition to college as normal and nondiagnostic of their belonging—that is, that uncertainty about belonging is “common and temporary” (Walton & Cohen, 2007, 2011; Wilson & Linville, 1982; Yeager, Walton, et al., 2016). Again, a Lay Theory × Psychological Affordances interaction emerged among disadvantaged students (students of color and first-generation/low-income college students), such that the treatment effects were far larger in college contexts that were more welcoming to disadvantaged groups, relative to less-welcoming contexts. Walton and colleagues (2020) argued that students in contexts that were not welcoming to their group may have felt that the intervention’s message—that belonging concerns are “common and temporary”—did not ring true; thus, undermining its long-term effects.

Although these findings are consistent with our model, a key limitation of the NSLM and CTC studies is that contextual moderators were measured, not manipulated. Measurement of contexts is an informative and suggestive step in many cases, particularly when experimental manipulations of the context are not possible or when the goal is to study large, representative populations in which it would be challenging to randomize contexts. Yet, interpretation of results in these cases can be difficult because measured contextual variables may be confounded with many other differences across schools. In the NSLM, for instance, it is possible that schools with greater challenge-seeking norms also had more experienced teachers or higher quality curricula. Although these objective affordances might be critical for policy, they are not the same as the psychological affordances described in the model presented above. Likewise, because families are not randomly assigned to neighborhoods, schools with stronger challenge-seeking norms may be populated with students from families better able to benefit from the intervention. There is no strong reason to draw these alternative conclusions about the NSLM and CTC results, but the possibility itself is a justification for a fully experimental test of affordances.

Further, although previous studies are consistent with moderation by affordances in a general sense (e.g., the overall school culture relevant to challenge-seeking or belonging), they do not say much, if anything, about the aspects of the contexts that provide support for the proffered lay theory (e.g., the specific actions that an educator could take to provide

affordances). An important question remains: what situational cues could students be receiving that make the soil more fertile for the effects of a lay theory intervention?

The Present Research

Here we report a causal test of the hypothesized Lay Theory \times Psychological Affordance interaction. We conducted a field experiment in an English class in an urban middle school in which underperformance among language minority students (i.e., students who were not native English speakers) was a stated concern. Our study consisted of two experimental manipulations: a student lay theory intervention (delivered online) and a manipulation of the psychological affordance of the lay theory (delivered through a written note from their teacher). The primary dependent variable was students' performance on a foundational skill-building writing assignment, to which teachers' hand-written notes (randomly assigned) were appended.

The first manipulation was an online purpose lay theory intervention, which was updated and adapted from previous research for our study population (Yeager, Henderson, et al., 2014). A purpose lay theory has three mutually-supportive elements: a focus on learning and mastery, a self-oriented element ("this benefits me"), and a self-transcendent element ("this serves a purpose beyond me"). These three elements have been shown to be effective in counteracting the cultural lay theory dominant in most U.S. public schools, which is a more instrumental view of education (get good grades, get a job, and make money; see qualitative research by Yeager & Bundick, 2009). The purpose lay theory emphasizes these elements in the context of difficulty in school, particularly tedious or boring coursework, because lay theory interventions in general are most effective when they help people reinterpret causally ambiguous adverse experiences (see Ferrer & Cohen, 2019).

The second manipulation involved the purposeful affordances of the academic context, via a message from students' teachers. We used a method developed in previous research on psychological interventions among middle school students: the *written note paradigm* (Yeager, Purdie-Vaughns, et al., 2014). In a past study of "wise feedback," students received a randomly assigned note hand-written by their teacher, appended to the critical feedback that the teacher provided on students' essays (Yeager, Purdie-Vaughns, et al., 2014). In the written note paradigm, it is possible to manipulate the content of the note at the level of the student and to keep teachers unaware of students' condition (as we explain below). This method, therefore, allows for manipulation of specific, actionable aspects of the academic context while holding other factors of the class constant. The written note paradigm also provides greater statistical power because of student-level (rather than cluster-level) random assignment.

The dual treatment condition, therefore, provides students with a lay theory that is then echoed in their later experience: Students first receive a message that underscores the purposes behind learning and later have a specific exchange with their teacher that reaffirms that message. Why did the hand-written note come from teachers and not some other source? Teachers are a powerful medium to create psychological affordances. They are the leaders of the classroom; they set the agenda of the class and provide credible information

related to the purpose of learning. They are the authority in establishing norms, and they are in charge of the incentive structures that affect student compliance. Teachers are, in effect, the arbiters of meaning for classroom experiences. By contrast, other school staff members, such as nonteachers and administrators, are not in as direct a role to shape the meaning of classroom events.

We answer these questions in the context of an applied problem: performance in an English class among racial, ethnic, and language-minority (nonnative English-speaking) middle school students in a low-performing middle school. Past studies of purpose have focused on math and science performance among lower-achieving students in general, and included high school or college students (Paunesku et al., 2015; Yeager, Henderson, et al., 2014). Purpose has not yet been studied among underrepresented groups in middle school English/Language Arts. We selected this subject for the present study because our school partner was interested in the writing performance of the nonnative English-speaking population (that made up the majority of the student body). We explored whether the interaction between the manipulations appeared more strongly or exclusively among the students who were in the group of nonnative English speakers. In doing this exploratory analysis, we limited ourselves to only one moderator, and we used a conservative, Bayesian machine-learning algorithm that included penalties (i.e., strong prior distributions) to discourage the model from finding spurious results. Materials, data, and syntax are posted here: <https://osf.io/3ktsb/files/>.

Method Participants and School Context

Participants were $N = 321$, 7th and 8th grade students attending an urban, public middle school in one of the largest and most heavily populated urban areas in the United States.¹ We invited all 7th and 8th grade students in the school to participate in the study. Our sample size and the resulting statistical power were determined by the constraints of the school partner and could not be affected by our knowledge of the results.² Hence, sample size was not a degree of freedom that we could exercise post hoc.

The school site typically receives low ratings for its performance: fewer than 25% of its students meet the state's minimum threshold for grade-level proficiency in reading and writing (earning it a four out of 10 on www.greatschools.org in the year of the study). Nevertheless, the school had the capacity to support student success. It recently hired a new and progressive principal, well-trained and long-tenured teachers worked there, and its leadership had cultivated robust connections to prominent national professional development networks. In private interviews, all of the English teachers in the study expressed a strong interest in supporting students' success and a belief that their students could learn.³ Therefore, this school provided a good location for a first test of our hypotheses, because

¹This number represents the total sample of students for whom we have at least partial data for both the Time 1 and Time 2 assessments. The specific n for some of the analyses reported below was less than 321 because of missing data. Missingness did not vary by condition.

²Although we could not control our exact sample size, we chose to work with this school site in part because it had a large enough student body to provide reasonable statistical power given our research questions.

³These structural factors may be a necessary precondition for the effects of lay theory interventions and the cues that afford them to unfold. For instance, in schools with poor-quality curriculum or ineffective teachers, students may be unable to benefit from skill-building opportunities, even with adaptive lay theories and supporting psychological affordances. Further, if the treatment notes felt completely inauthentic, teens would have noticed.

students were not already optimally performing, but the messages from the lay theory intervention had the potential to take root.

Participants' ages were 12 (19%), 13 (43%), 14 (32%), or 15 (5%) years old. Forty-seven percent identified as female and 53 identified as male. The majority of participants came from groups that are underrepresented or disadvantaged in advanced educational contexts. In terms of social class, the majority of our sample came from low socioeconomic status (only 15% of mothers had a 4-year college degree or higher, one marker of higher socioeconomic status; Hauser, 1994). In terms of race/ethnicity, 58% were Hispanic/Latinx, 14% Black/African American, 17% Asian/Asian American, 7% Multiple racial/ethnic identification, and 2% White, non-Hispanic. In terms of linguistic minority status, 71% reported that they were nonnative English speakers. This high rate of nonnative English speakers is one reason why the school worked with us: they hoped that our intervention could help nonnative students who were struggling in English/Language Arts.

Procedure

Overview.—The study featured a 2 (student purpose intervention vs. control) \times 2 (purpose-affording note vs. control note) between-subjects design. On an exploratory basis, we examined how the effects of the experimental manipulations differed by nonnative English-speaking status. We aimed to recruit a diverse sample with as many students as possible. Therefore, we obtained parental consent for students to participate in the study (that was considered a “program evaluation” by the institutional review board at our institution) earlier in the year through a passive consent form in accordance with section §99.31 (a)(6)(i) of the Family Education Right to Privacy Act (FERPA). Students were included only if they had actively assented in the context of the baseline survey. We conducted analyses for scientific purposes, designated as “secondary analysis” (as defined by our institutional review board) of the program evaluation dataset, after the main results of the intervention had been communicated back to the school and the school staff through professional development workshops. This procedure yielded a high consent rate that allowed the study to include a relatively large percentage of students from socially disadvantaged families, who might normally be underrepresented in evaluations of educational programs that require active parental consent. This arrangement with the school necessitated our decision to use a control group that would be positive (i.e., one that emphasized learning), so that all students would benefit. This means the study is not designed to provide an estimate of the effect of the purpose lay theory on its own.

The study consisted of three phrases. First, students completed an online survey in January (Time 1) in the school's computer lab during their regular English class meeting time. During the Time 1 survey, students first answered baseline and demographic questions. They were then randomly assigned, at the student level by the survey software, to complete either the online purpose intervention or the control activity.

Two weeks later (in February), all of the 7th and 8th grade English teachers hand-wrote stacks of notes that were later appended by researchers to the writing assignment that students would later be asked to complete. These notes were randomly assigned, with randomization blocked by teacher, gender, and intervention condition strata (Time 2). Figure

1 presents an image of the assignment as experienced by students. The writing assignments were then distributed to students during their English class. To keep teachers blind to students' condition, the first page was a cover sheet printed with the student's name, which obscured from teacher's view the note students received. Students completed the assignment and then turned it in to their teachers. Teachers had been instructed by the researchers not to look at the completed assignments and to place all of them in a sealed envelope for delivery to the researchers. All complied with this request.

Researchers removed the notes from the assignments, again to keep teachers blind to condition, and then returned the assignments to the teachers. Teachers then graded the assignments (described below), recorded the scores, and sent the scores to the researchers. At this time, we did not collect additional survey measures because we did not want to raise suspicion among students that the note might be related to a research study or otherwise contaminate the experience of the experimental manipulations. Finally, 2 weeks after the note manipulation and 1 month after the baseline survey and purpose intervention, students returned to the computer lab to complete a follow-up survey (Time 3). During this survey session, students completed a number of measures tapping psychological changes that might underlie the degree of treatment efficacy (e.g., sense of purpose). These latter measures were collected on an exploratory basis; further details related to these measures are provided in the supplement.

Student purpose intervention.—The purpose for learning intervention has been tested in past studies (Paunesku et al., 2015; Yeager, Henderson, et al., 2014) and implements persuasive techniques now common in lay theory interventions in education. First, it presents students with survey statistics from their peers that establish a prosocial norm for learning. Second, to further reinforce this norm, the intervention showcases testimonials from peers who describe wanting to learn and build skills in school to open doors to a better future for themselves and others. Third, to facilitate internalization of the message, the intervention invites students to write persuasive letters to future students, inviting these students to keep their purpose for learning in mind when they feel bored, frustrated, or unmotivated (E. Aronson, 1999).

The control condition used the same three techniques described above (i.e., survey statistics, normative quotes, and letter to a future student); however, the content focused primarily on learning and developing academic skills. It did not connect learning in school to benefits for the self and others. The control group included a focus on learning and mastery—that previous research suggests may be motivating to students (Dweck & Leggett, 1988)—because we wanted to provide all students with content that could plausibly benefit them, for ethical reasons. The nature of the control condition, therefore, provides a conservative test of our hypotheses. To keep length consistent across the two conditions, the control condition also presented neutral, filler content related to how middle school is different from elementary school. This content was similar to the control conditions used in previous lay theory intervention studies (Goyer et al., 2019; Yeager, Henderson, et al., 2014).

Before administering the purpose for learning intervention, we adapted it for our sample of middle school students. As noted, the purpose for learning intervention used in prior

research was developed for older adolescents in high school and for adults in college, and so it was developmentally inappropriate. Furthermore, the present study was conducted in an English class, and so we included more content about why it is important to be able to read and communicate effectively in writing. To tailor the purpose intervention to our young adolescent sample, we used a year-long research and development (R&D) process (the process, applied to a different intervention, is described in Yeager, Romero, et al., 2016). See the online supplemental materials for details on this R&D process.

Skill-building writing assignments.—Working closely with teachers, we created two different writing assignments, one for each grade level used in our sample, to which our note manipulation would be appended (the full assignments are in the online supplemental materials). These assignments, which were based on guidelines for English language proficiency standards by grade level in the state of Texas, asked students to read a passage selected by their teacher, appropriate to grade reading levels, and then to write a short essay about the theme of the passage. Teachers in both grade levels confirmed that the assignments were consistent with the skill-building writing assignments included in their regular course curricula. They also confirmed that identifying and reflecting on the major themes in reading passages was a critical skill, but one that was difficult for many students to see the purpose. Thus, the assignment used in the present study provided a valid skill-building opportunity that could benefit from a shift in students’ mindset toward the purpose lay theory.

Note manipulation.—Across all conditions, the notes were hand-written by students’ teachers and administered as part of a regular classroom activity. Because the notes were hand-written, students were led to see them as a personal message from their teacher.

Table 1 presents the specific notes that were randomly assigned to students. The focal *purpose-affording note* highlights three elements of a purpose for learning: (a) a focus on skill-development, (b) a future benefit to the self from having stronger skills, and (c) a future benefit to others from having stronger skills. The note stated, “*I’m giving you this assignment because I think you have the potential to **get an interesting job [2]** and **make people’s lives better one day [3]**, if you **develop your skills [1]** on assignments like this one” (numbers and bold emphasis added to highlight the message’s invocation of the key elements of a sense of purpose).*

Half of students were randomly assigned to receive the purpose-affording note, and the other half were randomly assigned to receive one of three control notes. To enhance experimental precision, each of the control notes systematically removed one of the three elements in the purpose intervention. One of the control notes removed the self-transcendent component of the note but maintained a focus on skill improvement and its benefits to self (Control note 1; see Table 1). Another removed both the self-transcendent and self-benefit components but included the focus on improving skills (Control note 2). A final control note removed all three components and was entirely “placebic” in its explanation for the assignment (Control note 3; see Langer & Rodin, 1976). The primary analysis combined the three control groups because the power allocation decisions in our study were intended to optimize the contrast between the purpose-affording note and all others, and because the three control notes did not statistically differ from each other.⁴

Previous research suggests that providing *any* rationale for a task can be motivating (Steingut, Patall, & Trimble, 2017). Thus, the control groups might increase students' scores compared with baseline, potentially shrinking the treatment contrast. However, for reasons related both to experimental control and to ethical considerations, we chose to have all students receive some form of a rationale. That is, all received a message that, research suggests, could plausibly benefit them. To ensure that the notes were authentic in all cases, we asked all teachers if they believed each of the relevant messages. All affirmed that they did.

R&D to design the purpose-affording note.—To select the appropriate language for the purpose-affording note, we engaged in an extensive review of past research, user-centered design, and A/B testing process (i.e., simple randomized experiments testing minor variations of possible notes) to create an impactful and credible purpose-affording note for this age group (Yeager, Romero, et al., 2016). First, we conducted a research synthesis using common methods from meta-analysis to consolidate the literature related to any study in which a learner was randomly assigned to a purpose or rationale for a learning task (not reported here). We examined the rationales that showed the strongest and weakest effects, and attended to features such as gain-framing (that we ultimately used) versus loss framing.

Second, informed by this process, we wrote eight potential notes that varied along three dimensions: short-term benefit versus long-term benefit, benefits that are intrinsic to the learning experience (i.e., enjoyment) versus extrinsic (i.e., earning more money), and benefits that are intended for the self (i.e., get a good job) versus benefits that transcend the self (i.e., making a difference for others one day). We conducted the A/B testing of different notes by randomly assigning a previous cohort of $N = 150$ 8th grade students at the school where we would eventually be conducting the experiment. They took part in a simulation study in which they reviewed all the notes and imagined receiving them from their teachers. In the A/B test, the order of the notes was counterbalanced and the notes were presented after students had first rated a teacher who provided an assignment without a note. Students then imagined that a teacher had given them each of the presented notes and rated how much they would trust the teacher, how much the note supported their purposes, and how motivated they would be. Although not all contrasts were statistically significant, the overall conclusion of this test was that long-term, intrinsic-motivation-oriented, and self-transcendent notes were consistently rated the most motivating—and by far exceeded the “try hard so you can make money” or “school can be fun and interesting” alternatives. This low-stakes study gave us early evidence in support of the present hypotheses and informed the content and wording of the note manipulation used here.

Next, in the year that the main study was run, we conducted focus groups with a subsample of students from the participating middle school. The students in these focus groups did not

⁴On an exploratory basis, we conducted analyses of the self-oriented note relative to Control notes 2 and 3. Mirroring previous research on a purpose for learning (Yeager, Henderson, et al., 2014), the self-oriented note (Control note 1), which had a rationale but not a self-transcendent one, fell midway between the purpose-affording note and other two control notes. Because the self-oriented + skill-development note was a part of the control group, all effect sizes for the purpose-affording note reflect a conservative test of the hypothesis. Removing the self-oriented + skill-development note from the control condition, of course, strengthens all of the reported effect sizes for the purpose-affording note.

later participate in the experiment. We asked students to read a series of notes from teachers, presented in a packet, and to imagine, for each, that their teacher had given it to them. They reflected on the meaning of each one (see Figure 2 for examples). Our analyses were qualitative, based on both students' written responses and the content of their discussion. Note that in panel A, on the left, when students read a note claiming that an assignment would be "useful," they seemed to react negatively and with mistrust. One student, after reviewing the note, "I'm giving you this worksheet because I think it's interesting," stated, "That's what teachers say when they know it's *not* interesting." By contrast, the purpose-affording note, to our surprise, was met with near-universal support. One student read the note and said, "Finally! A reason for learning that actually matters. I would do whatever this teacher said." See another example in Figure 2, Panel B. Results from this focus group provided early encouraging evidence that the purpose-affording note resonated with students.

Measures

Baseline measures.—A number of psychological variables were assessed before random assignment, to provide baseline covariates and to assess the effectiveness of random assignment. The survey also asked students to report their year and month of birth (to calculate age) and their identified gender, in addition to variables listed below.

Baseline self-transcendent reasons for learning.—Students answered two questions to assess their self-transcendent reasons for learning: "The main reason why I learn in school is so I can make a difference for my community or family one day" and "The main reason why I learn in school is so I can make a difference for other people one day." The items were rated on a scale from 1 = *not at all true* to 5 = *extremely true*. The two items were correlated at $r = .56, p < .001$. Therefore, we combined them by taking their unweighted average, so that higher values corresponded to more self-transcendent reasons for learning.

Baseline school trust.—Previous research found that a different written note manipulation increased school trust for racial minority students (Yeager, Purdie-Vaughns, et al., 2014). Therefore, we assessed trust in the present study. Students answered four items assessing school trust, adapted from previous research on middle school students (Yeager, Purdie-Vaughns, et al., 2014; e.g., "I am treated fairly by teachers and other adults at [NAME] Middle School,"). Each was rated on a 6-point scale, 1 = *strongly disagree*, 6 = *strongly agree*. The responses were averaged into a composite, such that higher values reflected greater trust ($\alpha = .73$).

Baseline perceived support for purpose in English class.—To control for students' pretreatment experiences with their teachers, the baseline survey also asked students about whether teachers supported their purposes. Students rated how true these three statements were: "I am certain that my English Literature/Language Arts teacher cares about my personal goals and interests," "I am certain that my English Literature/Language Arts teacher believes I can do something important with my life," and "My English Literature/Language Arts teacher always explains why the class assignments are important." The items were rated on a scale from 1 = *not at all true* to 5 = *extremely true*, and then

averaged into a composite, with higher values indicating more perceived support for purpose ($\alpha = .61$).

Nonnative English speaker status.—To assess language minority or nonnative English speaker status, students reported on whether English was the first and main language they spoke with their families, 0 = *native English speaker* (29%), 1 = *not a native English speaker* (71%).

One-month follow-up measures.—Students completed a survey 1 month after the student purpose intervention that again assessed trust, purpose, and the personal meaningfulness of school overall. These measures were assessed on an exploratory basis to examine the psychological and motivational outcomes of the online intervention. As a kind of manipulation check, we expected the lay theory intervention to increase a sense of purpose a month later for all students, and this is what the results showed (see the online supplemental materials).

Performance on writing assignment.—The main dependent variable was students' performance on the writing assignment to which the hand-written note was appended. This assignment constituted a formative assessment and was graded by teachers in concordance with the Texas state standards for assessments of this type. Teachers were blind to student condition. They evaluated whether students' work met grade-level progress for their writing. They are trained by the district each year to use the state scoring rubric and so no additional training from researchers was needed. The state standards assessed by the writing activity encompassed: (a) an ability to compose text with a clear, controlling idea, (b) coherent organization, (c) sufficient development and effective use of language and conventions, (d) use of appropriate facts and details and no extraneous information or inconsistencies, (e) accurately synthesizes ideas from several sources, and (f) uses a variety of sentence structures and transitions to link paragraphs.

Teachers assigned each activity a score that ranged from 0 to 3, where 0 = *very limited writing performance (or did not do the assignment; 14% of students overall)*, 1 = *basic writing performance (34% of students)*, 2 = *satisfactory writing performance (34% of students)*, 3 = *accomplished writing performance (17% of students)*. According to the state, scores of 2 and 3 reflect adequate proficiency. Roughly 50% of students scored below satisfactory, consistent with the school's historical test records as reported by www.greatschools.net.

Analysis Plan

The study was conducted in 2015, when preregistration of analysis plans was still a relatively uncommon practice. We refer to analyses as “exploratory” when they were not documented in grant proposals or Institutional Review Board (IRB) documentation before obtaining the data. We also limited ourselves to two primary analyses and one main outcome to reduce the likelihood of false discoveries that come from testing many factors and outcomes.

All analyses were “intent-to-treat” (ITT), which means that participants were included in analyses provided that they were randomly assigned to both manipulations, regardless of fidelity of implementation. This is the most conservative analysis but is the best way to ensure a valid causal inference. Analyses were conducted with linear regression models. As a robustness check, we also conducted a series of mixed effects linear regression models with students nested within teachers and random intercepts (see online supplemental materials for more information about these analyses); the conclusions were the same because the intraclass correlations (ICCs) for students nested within teachers were very low.

To reduce error variance in the outcomes and to control for any chance differences between conditions, all analyses included covariates for teacher (coded as a series of dummy variables), the three baseline psychological variables, and the interactions among the three baseline psychological variables. A comprehensive covariate set is especially important given our interest in higher-order interactions (between two condition variables and the English-language status variable). Higher-order interactions split the sample into smaller and smaller groups, which could increase the chance of failures of random assignment within subgroups and, therefore, increase the need to control for baseline differences that are correlated with the outcome (a version of the “multivariate outlier” problem). Because choices of covariates are researcher degrees of freedom that can sometimes be exercised in an anti-conservative manner, we supplemented the ordinary least squared (OLS) regression model with a Bayesian robustness analysis that relied on a penalized machine-learning algorithm to flexibly incorporate covariates and moderators. Reassuringly, this showed that none of our conclusions depended on the choice of covariates. We also report the primary analyses without covariates and with only the main effects of the covariates in the online supplemental materials.

Here and throughout, all descriptive statistics (means and standard deviations) are based on the raw data. Null hypothesis tests derived from OLS regression models were calculated with the “Anova” function in the R package “car” (Fox & Weisberg, 2011). Average marginal effects (also called “simple effects” in much psychological research) were estimated using the “margins” package in R (Leeper, 2018), which implements one of the most typical methods for estimating marginal effects in models with higher-order interactions across several social science disciplines (sociology, economics, and political science). *p* values are based on two-sided tests, and standardized mean differences (*SMDs*) are equal to the average marginal effect divided by the control group’s raw standard deviation, following standard practice for educational evaluation studies. Finally, a Bayesian robustness analysis, which incorporates the uncertainty that comes from the possibility of detecting interaction effects by chance into its model output, was implemented using the “BCF” package in R (Hahn, Murray, & Carvalho, 2019).

Results

Research Question 1: Did the Purpose-Affording Note Moderate the Effects of the Purpose Lay Theory Intervention?

Consistent with our hypotheses, our primary analysis found a significant two-way Student Purpose Intervention × Purpose-Affording Note interaction on students’ assignment scores,

$F(1, 287) = 5.958, p = .015$. Breaking the two-way interaction down, there was a significant simple effect of the student purpose intervention when paired with the purpose-affording note. Among students who received the purpose-affording note, those who received the student purpose intervention earned higher scores on the writing assignment relative to those in the control condition, AME (i.e., average marginal effect, or simple effect) = 0.307, SE (standard error) = 0.152, $z = 2.028, p = .043, SMD$ (i.e., standard mean difference) = 0.330, 95% confidence interval, CI [0.011, 0.650]). However, when participants did *not* receive the purpose-affording note, there was no benefit of the student purpose intervention relative to the control, $AME = -0.194, SE = 0.144, z = -1.347, p = .178, SMD = -0.208, 95\% CI [-0.510, 0.095]$. Thus, consistent with the psychological affordance model of lay theory interventions proposed here, the student purpose intervention increased performance only when the treatment message was supported in the classroom context.

On an exploratory basis, we also examined the simple effects of the purpose-affording note. The purpose-affording note only had a positive effect among students who had earlier received the purpose intervention ($AME = 0.313, SE = 0.146, z = 2.151, p = .032, SMD = 0.336, 95\% CI [0.030, 0.643]$). A purpose-supportive note from a teacher did not have a significant effect among students who received the control activity ($AME = -0.188, SE = 0.148, z = -1.271, p = .204, SMD = -0.202, 95\% CI [-0.513, 0.109]$). This analysis suggests that students may have needed a broader cognitive context to make sense of and be motivated by the one-time note.

Research Question 2: Did the Manipulations Benefit Nonnative English Speakers to a Greater Extent?

As noted, psychological interventions in general tend to yield the strongest benefit among subgroups of students who are the most marginalized or underserved in education (J. Aronson, Fried, & Good, 2002; Walton, 2019; Walton & Cohen, 2007, 2011; Yeager, Walton, et al., 2016) because the interventions attempt to fulfill a psychological need that has systematically gone unmet for an individual or group. In the English class examined here, the relevant lower-achieving subgroup was nonnative English speakers. We suspected they might be the most alienated from the English course content (Gluszek & Dovidio, 2010), and that they might be the most likely to resonate with the self-transcendent values emphasized in the purpose manipulations. Native English speakers, by contrast, may not have needed the intervention (because they were already higher-achieving) and because the control condition might have already been sufficiently motivating (because it emphasized learning and mastery in a subject in which they do well).

Consistent with this logic, analyses revealed a three-way Student Purpose Intervention \times Purpose-Affording Note \times Nonnative English Speaker interaction, $F(1, 287) = 4.425, p = .036$. Table 2 displays the relevant means. Breaking down the three-way interaction, among nonnative English speakers (the numerical majority, $N = 225$), there was a strong and significant two-way Student Purpose Intervention \times Note Condition interaction, $b = 0.783, SE = 0.246, t(287) = 3.188, p = .002$. A strong positive effect of the online student purpose intervention emerged among those who also received the purpose-affording note, $AME = 0.546, SE = 0.177, z = 3.080, p = .002, SMD = 0.587, 95\% CI [0.213, 0.960]$. Among those

who did *not* receive the purpose-affording note, the effect of the purpose intervention was nonsignificant as expected, $AME = -0.237$, $SE = 0.172$, $z = -1.380$, $p = .168$, $SMD = -0.254$, 95% CI [-0.616, 0.107].

Within the group of native English speakers ($N = 91$), the two-way interaction and the simple effects within each note condition were all nonsignificant ($ps > 0.3$). That is, controls, who received positive messages about learning and mastery but not self-transcendent messages, performed just as well as treated students, in the subgroup of native English speakers.

As in the full sample, the converse moderation pattern also emerged. Among nonnative students, the purpose-affording note led to higher performance on the writing assignment among those who received the student purpose intervention ($AME = 0.476$, $SE = 0.169$, $z = 2.82$, $p = .005$, $SMD = 0.511$, 95% CI [0.156, 0.867]), but not among those who received the control activity ($AME = -0.307$, $SE = 0.178$, $z = -1.725$, $p = .085$, $SMD = -0.330$, 95% CI [-0.704, 0.045]). Once again, among native English speakers, the two-way interaction and simple effects of the purpose-affording note were all nonsignificant ($ps > 0.6$).

As a follow-up analysis, we assessed whether the student purpose intervention and purpose-affording note *combined* closed the performance gap between native and nonnative speakers. To do so, we created a variable that tested the combined effect of both experimental manipulations (1 = *online purpose intervention + purpose-supportive note*, 0 = *all other conditions combined*). Among students who did not receive the combined interventions (three-fourths of the data), there was a significant gap in performance such that nonnative speakers performed worse than their language majority peers on the writing task, $b = -0.302$, $SE = 0.133$, $t = -2.281$, $p = .023$. Among students who received the combined purpose interventions, however, there was no longer a significant achievement gap, $b = 0.297$, $SE = 0.241$, $t = 1.231$, $p = .219$, and indeed, if anything, the gap reversed. When students reflected on their purposes and when a teacher subsequently reinforced it, language minority students achieved performance that was statistically equivalent to their more advantaged peers.

Why did nonnative English-speaking students appear to benefit more from the student purpose intervention and purpose-affording note relative to their peers? One possibility is that the purpose-affording messages compensated in some way for the psychological environment that language-minority students desired but were not receiving. Indeed, in an exploratory analysis we found that nonnative English speakers endorsed more self-transcendent motives for learning on the baseline survey, relative to native English speakers, and this result approached statistical significance $b = 0.230$, $SE = 0.121$, $t(321) = 1.907$, $p = .058$. This finding provides tentative support for the hypothesis that the self-transcendent purpose was a stronger cultural match for nonnative English speakers (Covarrubias, Herrmann, & Fryberg, 2016; Markus & Conner, 2013; Markus & Kitayama, 1991; Stephens, Fryberg, Markus, Johnson, & Covarrubias, 2012). Additionally, nonnative English speakers reported *less* support from their English teachers for their purposes for learning, as assessed by the baseline survey, although this too was marginally significant, $b = -0.177$, $SE = 0.100$, $t(321) = -1.764$, $p = .079$. The joint intervention may have helped these students feel that

they were now getting the support for their purposes that they desired but had felt relatively absent. These results may also explain why native English speakers did not appear to benefit from the purpose manipulations. They reported relatively weaker self-transcendent motives for school and perceived that their teachers supported their purposes to a greater degree. The purpose manipulations may have been less motivating among English-speaking students or may have been redundant with the support they were already receiving from the context.

Main Effects

There were no significant main effects of either manipulation on student scores, student purpose intervention versus control, $F(1, 287) = 0.259, p = .611, SMD = -0.050, 95\% \text{ CI} [-.171, 0.271]$; purpose affording note versus control note, $F(1, 287) = 0.5948, p = .441, SMD = 0.089, 95\% \text{ CI} [-0.129, 0.307]$. This does not mean that the interventions were ineffective, only that they were no more effective than the control conditions except within the theoretically expected cells. This is unsurprising because (a) the online control condition was a strong, prolearning treatment that might be expected to have significant benefits of its own, unlike past purpose for learning interventions, which were compared with more neutral control groups, (b) all of the control notes either encouraged learning or offered a placebo rationale, either of which could increase motivation to learn (Steingut et al., 2017). In this way, the present study was designed to determine if the match between a lay theory intervention and a later psychological affordance affected motivation, not to maximize the independent effects of each treatment.

Bayesian Robustness Analysis

A supplementary analysis assessed the robustness of the three-way interaction between the two treatment conditions and native English status. This was important because unpredicted moderation results have sometimes appeared because of chance patterns in the data (Bloom & Michalopoulos, 2013), and choices about model specifications can sometimes result in spurious results. Experts in field experiment methodology have recommended that analysts implement machine-learning methods to automate the search for interaction terms (Gerber & Green, 2012). To do this, we implemented a machine-learning algorithm called “Bayesian Causal Forest” (or BCF) that has recently emerged as a leading method in the field (Hahn et al., 2019), and is useful here for a number of reasons. First, BCF is conservative because it uses a prior belief that a treatment effect is modest on average and is not moderated. This “shrinks” all treatment effects to zero and to be similar. Thus, only strong evidence will cause the model to conclude that there is moderation, which guards against false-positive results. Second, the prior belief of no moderation reduces the sensitivity of the results to low statistical power in the subgroup of native English-speakers, because the model would shrink subgroup treatment effects toward the overall treatment effect if there were truly no moderation. Third, BCF can flexibly incorporate covariates using Bayesian Additive Regression Trees (BART; Hahn et al., 2019). That means that it can allow covariates to take on nonlinear functional forms and to interact with each other, reducing the potential for idiosyncratic choices about inclusion of covariates in the model to bias the results.

The highly conservative BCF algorithm provided strong evidence in support of the moderation results obtained in the linear model presented above, as shown in Figure 3. It

estimated a posterior probability of .99 of a positive, nonzero Student Purpose Intervention \times Note Condition interaction within the subgroup of nonnative English speakers, and a .95 probability of a positive, nonzero effect of the student purpose intervention among students in this subgroup who received the purpose-affording note. This marks a substantial shift in probabilities from a conservative prior probability of .50, and comes from an analysis that builds uncertainty about model specification into the posterior distribution. In short, the data support strong updating of beliefs toward the conclusions we reported.

Next, BCF estimated a posterior probability of .48 (or hardly any shift from a prior probability of .50) for the Student Purpose \times Note Condition interaction within the subgroup of native English speakers. This means that even though the model expected the two-way interaction with this subgroup to match the overall two-way interaction, and even though the native English speakers provided a minority of the data (and, therefore, would need substantially different patterns to end up with a different posterior distribution), the model nevertheless detected different results for native versus nonnative English speakers.

Finally, in a sensitivity analysis that reran the BCF using theoretically irrelevant moderators (such as grade level) no such pattern of moderation appeared (no posterior probabilities of a difference $>.55$). To summarize, the three-way interaction results were robust to a conservative Bayesian method that detects moderation when it is present and does not when it is not present.

Discussion

Lay theory interventions are a promising means for addressing academic underperformance, but not enough is known about the conditions under which they can be used to solve social problems (Walton & Wilson, 2018; Yeager & Walton, 2011; see also Cohen & Sherman, 2014). Here we presented the first experimental evidence that whether lay theory interventions are successful or not depends in part on whether the beliefs instilled by the intervention are supported and reinforced by the psychological affordances of the context. These findings suggest that students will be more apt to put lay theories into practice when they perceive cues suggesting that the lay theory is legitimate and adaptive in that context.

Support for this conclusion comes from the finding that a purpose for learning lay theory intervention led to improved performance on a skill-building writing assignment only when students later received a hand-written note from their teacher that supported their purpose for learning. This finding comes from the first randomized trial that manipulated both the students' lay theories *and* the psychological affordances for putting the lay theory into practice in a middle school classroom.

Although our primary predictions concerned the Lay Theory \times Affordance interaction, one might have expected at least a small positive main effect of the student intervention given previous research (Paunesku et al., 2015; Yeager, Henderson, et al., 2014). However, as we noted earlier, the study was not designed to detect this. Because we wanted all students to receive a potentially positive intervention, the control condition for our student purpose intervention focused on learning and skill development, and such messages can increase

students' motivation (Dweck & Leggett, 1988).⁵ Therefore, the present study was not designed to be informative with respect to the main effect of the online purpose lay theory intervention.

Moderation by Nonnative English-Speaking Status

An exploratory analysis found that nonnative English speakers were particularly disadvantaged in this context and showed the greatest benefits of the manipulations. As mentioned, one of the primary motivations for our partnership with the participating school was their concern about underperformance in English and writing courses among nonnative English speakers. Although the three-way interaction came from an exploratory analysis (and subgroup analyses can sometimes be unreliable, see Bloom & Michalopoulos, 2013), it was robust. Indeed, the conservative, Bayesian machine-learning algorithm assigned a strong posterior probability of differential effects by English-speaking status.

The exploratory analysis fits with previous research on lay theory interventions in education, which have typically found greater benefits among students who are members of marginalized groups or who chronically underperform relative to their potential (e.g., J. Aronson et al., 2002; Walton, 2019; Walton & Cohen, 2007, 2011; Walton & Wilson, 2018; Yeager, Walton, et al., 2016). Our findings further qualify this pattern by showing that disadvantaged students *in supportive contexts* benefit the most from the intervention. The intersection of vulnerability and opportunity may be the psychological “sweet spot” for lay theory intervention effects to appear (also see a meta-analysis by Ferrer & Cohen, 2019).

Implications for Psychological Interventions

The present research advances the growing literature related to how to replicate psychological (or “wise”) intervention effects on the performance of disadvantaged students (see Walton & Wilson, 2018). A critical consideration when scaling interventions like the one tested here is whether and when the context (the classroom, the school, or even broader society) is aligned with the intervention's message. If a few words on a note from a teacher could have so much impact on the effect size of a purpose for learning intervention in the group of lower-performing students, then it stands to reason that larger forces, if they were consistently applied, may have an even greater moderating impact. These include, classroom practices, school rituals, and cultural belief systems. More generally, our results suggest that students, and perhaps particularly those who are disadvantaged in a school context, are attentive to the difference between “talking the talk” and “walking the walk.” Perhaps it is not simply the message that “school has purpose” that matters for students who are disadvantaged in a context; instead what might matter more is that the message truly reflects the hearts and minds of the powerful adults with whom they interact.

Our results also suggest that psychological affordances have more of an impact when underperforming students can put them into the context of a larger meaning system. Recall that the purpose-affording note led to increased performance for nonnative English speakers

⁵The research that led to the present study (Yeager, Henderson, et al., 2014) did not find significant effects of the purpose intervention relative to a prolearning control group, but only with respect to a neutral control group.

only when they previously received the student purpose intervention. Lay theories may provide the necessary cognitive background that helps students, and contextually disadvantaged students in particular, make sense of the affordances and see them as part of a larger whole.

One benefit of the interactive framework proposed here is that it can unite two different kinds of intervention traditions. One, focused on situation-general beliefs (e.g., implicit theories, mindsets, belonging uncertainty beliefs), has sought to change students' worldviews and assumed that people will apply them in specific situations. Another, focused on situation-specific judgments (e.g., appraisals) has provided people with alternative ways of construing a particular situation, like a test or a moment of critical feedback, with the hope that people would generalize their construals to future situations (Brady, Hard, & Gross, 2018; Jamieson, Mendes, Blackstock, & Schmader, 2010; Yeager, Purdie-Vaughns, et al., 2014). It will be exciting to continue to integrate these two approaches and uncover the optimal circumstances for positive recursive processes to be sustained.

Implications for Educational Practice

This research has several practical implications. The question of how to address underperformance in English class—particularly among students who are not native English speakers—is a major applied concern among educators (Perie, Grigg, & Donahue, 2005). Nonnative English speakers face multiple barriers to success in the American education system and these barriers may be heightened in an English class, where their language minority status may be particularly salient (Gluszek & Dovidio, 2010). Our work is the first to use a field-experimental method to provide a social-psychological lens on this important issue. Indeed, our research suggests that underperformance among nonnative English students may arise, in part, from the fact that schools are failing to meet the psychological needs of these students. This work also suggests that one cost-effective strategy is to use targeted contextual and student-level approaches like the ones tested here. At the same time, our results may also speak to the potential power of more systemic efforts to reform classrooms to be more inclusive of the values of groups that are underrepresented in higher education.

Nevertheless, it is important to ask: Why did *native* English speakers not benefit from the manipulation? One possibility is mundane: they were already scoring higher than nonnative English-speaking students, so they may have shown less of a need for an intervention, or the subgroup may simply have been too small to detect a two-way interaction. Another possibility is more related to the theory at hand: perhaps those students were already receiving purposeful affordances in their classrooms. Indeed, more advantaged students may be more likely to receive rationales from their teachers that emphasize self-determination and internalized motivation (Solomon, Battistich, & Hom, 1996). Future research might, therefore, sample contexts in which purpose-supportive rationales were more or less abundant, and examine whether the note paradigm only shifts behavior when it compensates for the context.

Limitations and Future Directions

The present research is the first to examine the effects of lay theory interventions among nonnative English speakers—who represent one of the fastest growing demographic groups in the U.S. school system (Musu-Gillette et al., 2017), but have nevertheless been underrepresented in research on educational interventions. Because our hypotheses about this subgroup were exploratory, however, we were unable to examine all of the mechanisms through which they showed stronger effects of the manipulations. One possibility, supported by our exploratory analyses of survey data, is that the self-transcendent purpose intervention resonated more strongly with nonnative English speakers. Indeed, nonnative students reported directionally stronger prosocial motives for school and less support for their purposes relative to their peers at baseline ($p < .10$). The combination of the purpose intervention and the purpose-affording note may have provided a greater “cultural match” for such students (Covarrubias et al., 2016; Markus & Conner, 2013; Markus & Kitayama, 1991; Stephens, Fryberg, et al., 2012; Stephens, Townsend, et al., 2012). This evidence is not definitive, however, and future research should assess alternative possibilities.

An additional limitation concerns the nature of our primary dependent variable. Our primary contribution is theoretical, and we cannot support claims about downstream performance (e.g., grades or test scores). Whether the purpose lay theory intervention continues to affect performance over time likely depends on whether students continue to receive cues that the purpose lay theory is afforded in the classroom context. A fascinating direction of future research is to assess how more global changes to the psychological affordances in the classroom might yield larger performance effects that extend beyond a single assignment.

Conclusion

We started with the presumption that lay theory interventions can have effects on behavior, but may be heterogeneous: they may work better for groups who are in need of the intervention, and in contexts that afford the opportunity to act on the lay theory (see Walton & Yeager, 2020). This expectation was supported by the data we presented here. Nevertheless, there is more to understand about the interaction between individual beliefs and social contexts, and how to change both in beneficial ways. Only by uncovering these processes will we know the full potential of lay theory interventions to address important policy challenges for large populations of individuals.

This work was supported by the Raikes Foundation, the New Paths to Purpose Project at the University of Chicago Booth School of Business, the William T. Grant Foundation, the National Science Foundation under Grant HRD 1761179, by the National Institute of Child Health and Human Development (Grant 10.13039/100000071 R01HD084772–01, Grant P2C-HD042849), to the Population Research Center [PRC] at The University of Texas at Austin, a fellowship from the Center for Advanced Study in the Behavioral Sciences (CASBS), and a National Science Foundation SBE Postdoctoral Research Fellowship awarded to Stephanie L. Reeves (Award 1911643). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health and the National Science Foundation. We are grateful to Jazmine Kjar, Consuela Wright, and Cintia Hinojosa for assistance with data collection. We would also like to thank

the teachers and administrators of the school district, as well as our student participants and their parents, for their involvement in the project. Materials, data, and syntax are posted here: <https://osf.io/3ktsb/files/>.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Aronson E. (1999). The power of self-persuasion. *American Psychologist*, 54, 875–884. 10.1037/h0088188
- Aronson J, Fried CB, & 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. 10.1006/jesp.2001.1491
- Bates D, Mächler M, Bolker B, & Walker S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. 10.18637/jss.v067.i01
- Blackwell LS, Trzesniewski KH, & Dweck CS (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, 78, 246–263. 10.1111/j.1467-8624.2007.00995.x [PubMed: 17328703]
- Bloom HS, & Michalopoulos C. (2013). When is the story in the subgroups? Strategies for interpreting and reporting intervention effects for subgroups. *Prevention Science*, 14, 179–188. 10.1007/s11121-010-0198-x [PubMed: 21279547]
- Brady ST, Hard BM, & Gross JJ (2018). Reappraising test anxiety increases academic performance of first-year college students. *Journal of Educational Psychology*, 110, 395–406. 10.1037/edu0000219
- Cohen GL, Garcia J, Apfel N, & Master A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, 313, 1307–1310. 10.1126/science.1128317 [PubMed: 16946074]
- Cohen GL, Garcia J, & Goyer JP (2017). Turning point: Targeted, tailored, and timely psychological intervention In Elliot AJ, Dweck CS, & Yeager DS (Eds.), *Handbook of competence and motivation: Theory and application* (pp. 657–686). New York, NY: The Guilford Press.
- Cohen GL, 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. 10.1126/science.1170769 [PubMed: 19372432]
- Cohen GL, & Sherman DK (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology*, 65, 333–371. 10.1146/annurev-psych-010213-115137
- Covarrubias R, Herrmann SD, & Fryberg SA (2016). Affirming the interdependent self: Implications for Latino student performance. *Basic and Applied Social Psychology*, 38, 47–57. 10.1080/01973533.2015.1129609
- Destin M, & Oyserman D. (2010). Incentivizing education: Seeing schoolwork as an investment, not a chore. *Journal of Experimental Social Psychology*, 46, 846–849. 10.1016/j.jesp.2010.04.004 [PubMed: 20824201]
- Dweck CS, & Leggett EL (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256–273. 10.1037/0033-295X.95.2.256
- Ferrer RA, & Cohen GL (2019). Reconceptualizing self-affirmation with the trigger and channel framework: Lessons from the health domain. *Personality and Social Psychology Review*, 23, 285–304. 10.1177/1088868318797036 [PubMed: 30295141]
- Fox J, & Weisberg S. (2011). *An R companion to applied regression*. Thousand Oaks, CA: SAGE Publications.
- Gelman SA (2003). *The essential child: Origins of essentialism in everyday thought*. New York, NY: Oxford University Press 10.1093/acprof:oso/9780195154061.001.0001

- Gerber AS, & Green DP (2012). *Field experiments: Design, analysis, and interpretation* (1st ed.). New York, NY: Norton.
- Gibson JJ (1977). *The theory of affordances*. Hillsdale, NJ: Erlbaum.
- Gluszek A, & Dovidio JF (2010). The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review*, 14, 214–237. 10.1177/1088868309359288 [PubMed: 20220208]
- Gopnik A, & Meltzoff AN (1997). *Words, thoughts, and theories*. Cambridge, MA: The MIT Press.
- Goyer JP, Cohen GL, Cook JE, Master A, Apfel N, Lee W, ... Walton GM. (2019). Targeted identity-safety interventions cause lasting reductions in discipline citations among negatively stereotyped boys. *Journal of Personality and Social Psychology*, 117, 229–259. 10.1037/pspa0000152 [PubMed: 30920278]
- Hahn PR, Murray JS, & Carvalho C. (2019). Bayesian regression tree models for causal inference: Regularization, confounding, and heterogeneous effects. ArXiv:1706.09523 [Stat]. Retrieved from <http://arxiv.org/abs/1706.09523>
- Harackiewicz JM, Canning EA, Tibbetts Y, Priniski SJ, & Hyde JS (2016). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *Journal of Personality and Social Psychology*, 111, 745–765. 10.1037/pspp0000075 [PubMed: 26524001]
- Hauser RM (1994). Measuring socioeconomic status in studies of child development. *Child Development*, 65, 1541–1545. [PubMed: 7859541]
- Helms SW, Choukas-Bradley S, Widman L, Giletta M, Cohen GL, & Prinstein MJ (2014). Adolescents misperceive and are influenced by high-status peers' health risk, deviant, and adaptive behavior. *Developmental Psychology*, 50, 2697–2714. 10.1037/a0038178 [PubMed: 25365121]
- Hulleman CS, Godes O, Hendricks BL, & Harackiewicz JM (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology*, 102, 880–895. 10.1037/a0019506
- Hulleman CS, & Harackiewicz JM (2009). Promoting interest and performance in high school science classes. *Science*, 326, 1410–1412. 10.1126/science.1177067 [PubMed: 19965759]
- Jamieson JP, Mendes WB, 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. 10.1016/j.jesp.2009.08.015 [PubMed: 20161454]
- Langer EJ, & Rodin J. (1976). The effects of choice and enhanced personal responsibility for the aged: A field experiment in an institutional setting. *Journal of Personality and Social Psychology*, 34, 191–198. 10.1037/0022-3514.34.2.191 [PubMed: 1011073]
- Leeper TJ (2018). Margins: Marginal effects for model objects. Retrieved from <https://thomasleeper.com/margins/>
- Markus HR, & Conner A. (2013). *Clash!: How to thrive in a multicultural world*. New York, NY: Penguin.
- Markus HR, & Kitayama S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224–253. 10.1037/0033-295X.98.2.224
- Miller DT (1999). The norm of self-interest. *American Psychologist*, 54, 1053–1060. 10.1037/0003-066X.54.12.1053
- Miller DT, Dannals JE, & Zlatev JJ (2017). Behavioral processes in long-lag intervention studies. *Perspectives on Psychological Science*, 12, 454–467. 10.1177/1745691616681645 [PubMed: 28544860]
- Miller DT, & McFarland C. (1987). Pluralistic ignorance: When similarity is interpreted as dissimilarity. *Journal of Personality and Social Psychology*, 53, 298–305. 10.1037/0022-3514.53.2.298
- Miller DT, & Prentice DA (2016). Changing norms to change behavior. *Annual Review of Psychology*, 67, 339–361. 10.1146/annurev-psych-010814-015013
- Molden DC, & Dweck CS (2006). Finding “meaning” in psychology: A lay theories approach to self-regulation, social perception, and social development. *American Psychologist*, 61, 192–203. 10.1037/0003-066X.61.3.192

- Musu-Gillette L, De Brey C, McFarland J, Hussar W, Sonnenberg W, & Wilkinson-Flicker S. (2017). Status and trends in the education of racial and ethnic groups 2017. Nces 2017–051. Jessup, MD: National Center for Education Statistics.
- Paunesku D, Walton GM, Romero C, Smith EN, Yeager DS, & Dweck CS (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science*, 26, 784–793. 10.1177/0956797615571017 [PubMed: 25862544]
- Perie M, Grigg W, & Donahue P. (2005). The Nation's Report Card[TM]: Reading, 2005. NCES 2006–451. Retrieved from <https://eric.ed.gov/?id=ED486463>
- Petty RE, & Cacioppo JT (1986). The elaboration likelihood model of persuasion In Berkowitz L. (Eds.), *Advances in experimental social psychology* (Vol. 19, pp. 123–205). New York, NY: Academic Press.
- Schwartz DL, & Martin T. (2004). Inventing to prepare for learning: The hidden efficiency of original student production in statistics instruction. *Cognition and Instruction*, 22, 129–184. 10.1207/s1532690xci2202_1
- Solomon D, Battistich V, & Hom A. (1996). Teacher beliefs and practices in schools serving communities that differ in socioeconomic level. *Journal of Experimental Education*, 64, 327–347. 10.1080/00220973.1996.10806602
- Steger MF, Frazier P, Oishi S, & Kaler M. (2006). The meaning in life questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology*, 53, 80–93. 10.1037/0022-0167.53.1.80
- Steingut RR, Patall EA, & Trimble SS (2017). The effect of rationale provision on motivation and performance outcomes: A meta-analysis. *Motivation Science*, 3, 19–50. 10.1037/mot0000039
- Stephens NM, Fryberg SA, Markus HR, Johnson CS, & Covarrubias R. (2012). Unseen disadvantage: How American universities' focus on independence undermines the academic performance of first-generation college students. *Journal of Personality and Social Psychology*, 102, 1178–1197. 10.1037/a0027143 [PubMed: 22390227]
- Stephens NM, Townsend SS, Markus HR, & Phillips LT (2012). A cultural mismatch: Independent cultural norms produce greater increases in cortisol and more negative emotions among first-generation college students. *Journal of Experimental Social Psychology*, 48, 1389–1393.
- Trope Y, & Liberman N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117, 440–463. 10.1037/a0018963 [PubMed: 20438233]
- Vallacher RR, & Wegner DM (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, 94, 3–15. 10.1037/0033-295X.94.1.3
- Walton GM (2019). How do the effects of a brief social belonging intervention on college achievement vary across students and schools? Washington, DC: The Society for Research on Educational Effectiveness.
- Walton GM, & Cohen GL (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, 92, 82–96. 10.1037/0022-3514.92.1.82 [PubMed: 17201544]
- Walton GM, & Cohen GL (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science*, 331, 1447–1451. 10.1126/science.1198364 [PubMed: 21415354]
- Walton GM, & Wilson TD (2018). Wise interventions: Psychological remedies for social and personal problems. *Psychological Review*, 125, 617–655. 10.1037/rev0000115 [PubMed: 30299141]
- Walton GM, & Yeager DS (2020). Seed and soil: Psychological affordances in contexts help to explain where wise interventions succeed or fail. *Current Directions in Psychological Science*, 29, 219–226. 10.1177/0963721420904453 [PubMed: 32719574]
- Wilson TD, & Linville PW (1982). Improving the academic performance of college freshmen: Attribution therapy revisited. *Journal of Personality and Social Psychology*, 42, 367–376. 10.1037/0022-3514.42.2.367
- Yeager DS, & Bundick MJ (2009). The role of purposeful work goals in promoting meaning in life and in schoolwork during adolescence. *Journal of Adolescent Research*, 24, 423–452. 10.1177/0743558409336749

- Yeager DS, Hanselman P, Walton GM, Murray JS, Crosnoe R, Muller C, ... Dweck CS. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, 573, 364–369. 10.1038/s41586-019-1466-y [PubMed: 31391586]
- Yeager DS, Henderson MD, Paunesku D, Walton GM, D’Mello S, Spitzer BJ, & Duckworth AL (2014). Boring but important: A self-transcendent purpose for learning fosters academic self-regulation. *Journal of Personality and Social Psychology*, 107, 559–580. 10.1037/a0037637 [PubMed: 25222648]
- Yeager DS, Purdie-Vaughns V, Garcia J, Apfel N, Brzustoski P, Master A, ... Cohen GL. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General*, 143, 804–824. 10.1037/a0033906 [PubMed: 23937186]
- Yeager DS, Romero C, Paunesku D, Hulleman CS, Schneider B, Hinojosa C,... Dweck CS. (2016). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of Educational Psychology*, 108, 374–391. 10.1037/edu0000098 [PubMed: 27524832]
- Yeager DS, & Walton GM (2011). Social-psychological interventions in education: They’re not magic. *Review of Educational Research*, 81, 267–301. 10.3102/0034654311405999
- Yeager DS, Walton GM, Brady ST, Akcinar EN, Paunesku D, Keane L,... Dweck CS. (2016). Teaching a lay theory before college narrows achievement gaps at scale. *Proceedings of the National Academy of Sciences of the United States of America*, 113, E3341–E3348. 10.1073/pnas.1524360113 [PubMed: 27247409]

Student Name: _____
 Class Period: _____
 Date: _____

A Community Changes

Instructions:

- Read and annotate the text
- Circle new vocabulary words
- Define any new vocabulary words in the margin either by inferring it from the text or using a dictionary

Define any new vocab words here:

Pilsen is an old community in Chicago with a long history. This part of Chicago started small and got bigger, as most neighborhoods in the city did. Its cultural history is about moving, changing, and connecting. Many people have moved there over the decades. This movement started when in this part of the city when Chicago was growing rapidly. It was a person who moves from one country to another. Pilsen was first populated by German. They had moved from Germany. If you go to some of the places they built. But when they first moved to Pilsen, they found homes. Then immigrants continued to migrate there, Pilsen so they could be near people who came from other countries.

Many immigrants found homes and jobs in Pilsen, and they had to face a challenge and change for them. They had to figure out how to live in a new country.



and the whole community.

The newcomers built churches, created gardens, opened stores, and set up newspapers. The newspapers were published in their own language, and they were a way to educate the newcomers about their own community and the city through articles and editorials. Many entrepreneurs opened businesses, and soon the community was bustling. Some social workers supported this progress, they guided the immigrants who kept coming to the community. They set up settlement houses, where people could learn English and get help finding work and housing. But then there was movement out of the community, it was a local migration to suburbs and other parts of the city. People were looking for more opportunities. As the population declined, businesses closed.

Hand-written note

home in their new country, it was not easy to accomplish the changes. Cooperation and collaboration were important for the progress of each family

Short Answer Response

Read the question below. Answer the question and provide text evidence to support your thinking. Please PLAN your response before writing your final draft. You may use only the space provided (10 lines). Any text written outside of the box **WILL NOT** be considered nor graded as a part of your answer.

How do immigrant groups change the community that they live in? Use examples from the **first** group of immigrants in Pilsen and from the **later** group of immigrants in Pilsen.

Writing Plan	Explain your answer in your own words	Connect to text (copy text evidence here)
How did the first group of immigrants to Pilsen change the community?		
How did the second group of immigrants to Pilsen change the community?		

Figure 1.

The setup of the written note paradigm. Figure 1 depicts the skill-building writing assignment that 7th grade students were asked to complete, with randomly assigned appended note (i.e., the purposeful affordance manipulation) from teacher; the 8th grade assignment was parallel but called for an analysis of a poem. Grade level did not moderate the effects and, as such, we do not think the specific topic of the writing assignment affected the results.

A Imagine your teacher gives you a writing worksheet with this note attached to it:

I'm giving you this worksheet because I think it's useful.

What do you think a typical middle school student would think? Please write a few sentences about what you think a normal middle school student would think if they saw this note.

I think that if a student (normal middle school student) sees this, that he would probably think in his head: useful for what? Because, normal middle school students don't think that deeply, they don't even have an idea about their skills, and future and everything. Maybe the teachers should change the way they give work.

B Imagine your teacher gives you a writing worksheet with this note attached to it:

I'm giving you this worksheet because I believe you can get an interesting job and help people one day, if you develop your writing skills.

What do you think a typical middle school student would think? Please write a few sentences about what you think a normal middle school student would think if they saw this note.

This is an awesome note for the assignment that teachers given, it makes students to think about their future life and make them do it more confidently.

Figure 2. Examples of students reacting to notes during a pilot study focus group. Students in a pilot focus group (before conducting the present study) read through and commented on each note privately before discussing with the group; these written responses were, therefore, uncontaminated by discussion with peers or researchers.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

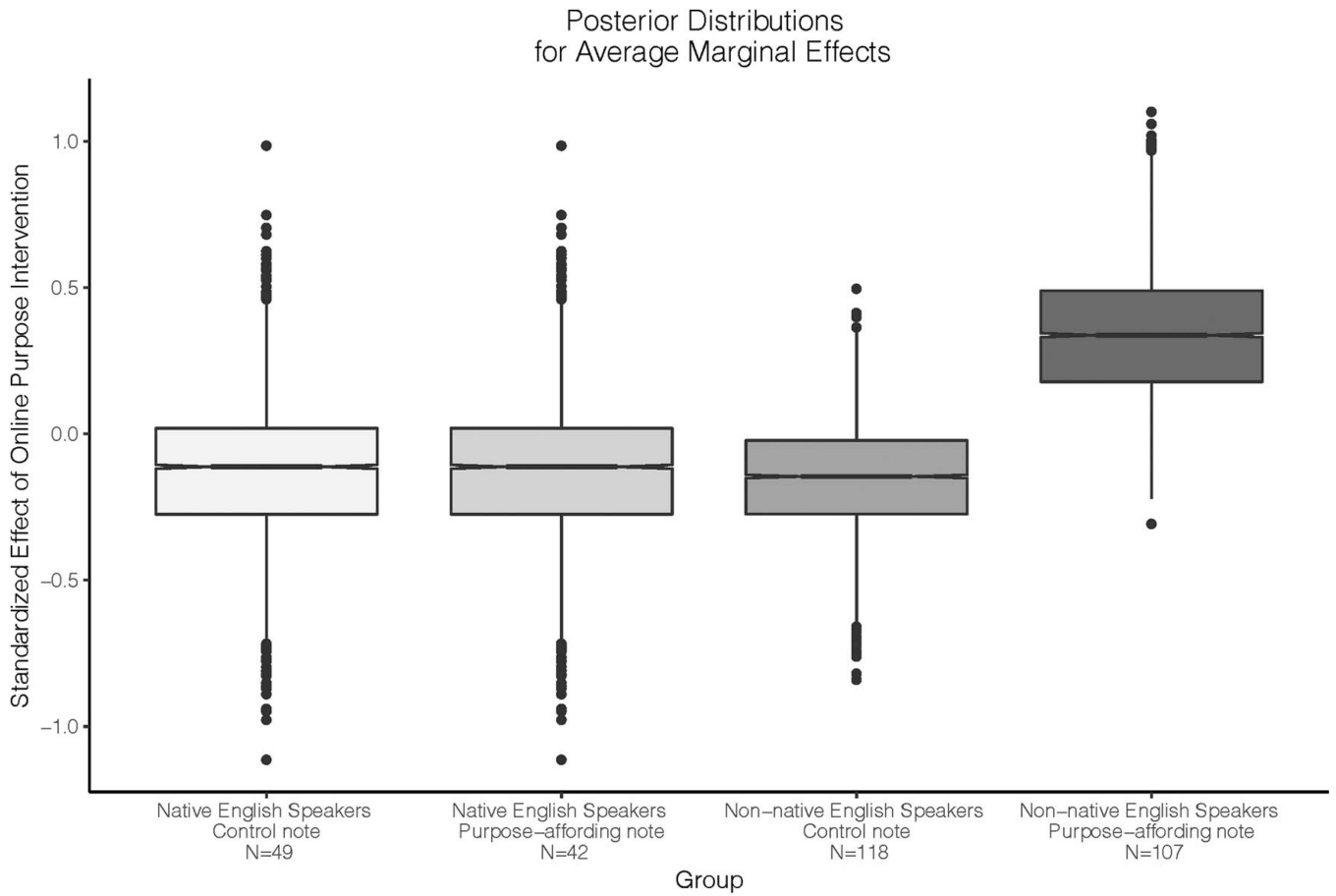


Figure 3. Treatment effects of the student purpose lay theory intervention on writing assignment scores (vs. the control condition) by note condition and nonnative English speaker status, estimated in a Bayesian Causal Forest analysis. The figure plots values drawn from the posterior distribution of individual treatment effects—one draw per participant. Posterior distributions are updated from a prior distribution centered at an effect size of 0 (the middle of the y -axis); therefore, the distances of the distributions from 0 represent the extent to which the data updated that prior belief of no effect (and no difference in effects across groups). Boxes represent the interquartile range, lines represent the middle 95% of the distribution, and individual data points are draws from the posterior that extend beyond the middle 95% of the distribution. The outcome ranged from 0 to 3 with a standard deviation of about 1, so the unstandardized effects on the y -axis are approximately the same as the corresponding standardized mean differences.

Table 1

The Text of the Purpose-Affording and Control Notes

Note condition	Text written by teacher on 7th to 8th grade students' assignments
Purpose-affording note (50% allocation)	
Self-transcendent + Self-oriented + Skill-improvement	"I'm giving you this assignment because I think you have the potential to get an interesting job and make people's lives better one day, if you develop your skills on assignments like this one."
Control notes (50% allocation)	
Control Note 1: Self-oriented + Skill-improvement	"I'm giving you this assignment because I think you have the potential to get an interesting job one day, if you develop your skills on assignments like this
Control Note 2: Skill-improvement	"I'm giving you this worksheet because I think it can help you practice your skills." ^a
Control Note 3: Noninformative	"I'm giving you this worksheet because it's the next activity in the lesson plan."

Note.

^aIn control Note 2, we changed the word "develop" to "practice" because, during piloting, we learned that adolescents interpreted "develop" in a pejorative way if it was not accompanied by a rationale, because of the implication that students lacked skills if they had to develop them. This was a necessary change to prevent any of the treatments from having a negative effect.

Table 2. Raw Means and Standard Deviations for Each Experimental Condition by Nonnative English-Speaking Status

Student group	Teacher note condition	Student lay theory condition					
		Control condition		Purpose intervention			
		M	SD	n	M	SD	n
Nonnative English Speakers	Purpose-affording note	1.31	0.87	51	1.80	0.89	56
	Control note	1.60	0.97	56	1.36	0.93	62
Native English speakers	Purpose-affording note	1.86	0.94	22	1.53	1.02	20
	Control note	1.74	0.86	27	1.67	0.80	22