tapraid5/zcz-edu/zcz-edu/zcz00316/zcz2768d16z | xppws | S=1 | 3/22/16 | 8:14 | Art: 2016-1329

Journal of Educational Psychology
2016, Vol. 108, No. 3, 397–404

0022-0663/16/$12.00 http://dx.doi.org/10.1037/edu0000122

COMMENTARY

The Half Empty Question for Socio-Cognitive Interventions

Daniel L. Schwartz, Katherine M. Cheng, Shima Salehi, and Carl Wieman
Stanford University

The studies in this special section of the *Journal of Educational Psychology* present a variety of social–psychological interventions across large numbers of classrooms and populations. They show notable benefits for many students at risk for low performance. This is the glass half-full interpretation, and we consider the strengths of the articles from this vantage. When viewed collectively, however, the results also raise a number of basic questions, the most pressing of which is, why do the interventions only benefit a particular population, when there are reasons to believe most students might benefit? This is the glass half-empty interpretation. We consider the limitations of each study from this vantage. We argue that although these studies provide important contributions, in all cases better measures and theories are necessary. We identify two related questions for the field: Why do these interventions show selective effects and fail to generalize to all students? Second, why do these very different interventions, presumably involving different psychological mechanisms, consistently result in similar selective effects? To answer these questions, a grand challenge for the field is to develop better instrumentation that can capture relevant behaviors and attitudes over time, and how these vary across context and population. With improved instrumentation, efficacy studies could rest on a much stronger foundation that would yield more confidence as to the broad benefits and replicability of large-scale social–cognitive interventions.

Keywords: motivation, socio-cognitive interventions, methodologies, selective effects

The central question of motivation research is why people take some goal-directed actions instead of others, including no action at all. It is a big question, perhaps intractable. Yet, it is ever present. To bring the question to size, the current authors of this special section of the *Journal of Educational Psychology* ask whether certain attitudes can put students on a trajectory to improved academic outcomes—outcomes the students would presumably want to achieve. The studies occur in classroom contexts, and for the most part, they involve attempts to change student beliefs. The authors are working on students' cognitions about their social circumstances in school; for example, "If I fail, what does it say about me as a person?"

There are other drivers of motivation besides socio-cognitive beliefs. The researchers might have investigated how rewards shape academic behavior, or they might have directly changed the participation structures of the classrooms themselves. The practical bet of the current work, however, is that it is possible to make brief interventions that change student beliefs and attitudes. These interventions will help the students see themselves in a light that encourages intellectual growth and at a much lower cost than changing the environment of schooling itself. It is important to emphasize that there is nothing in this approach that diminishes the importance of creating hospitable opportunities for intellectual and emotional growth. Socio-cognitive interventions are not a vaccine against a germ filled environment. They are one part of the complex story for why people take some goal-directed actions instead of others.

To appreciate the complexity of the story, one may note that even within the confined investigation of socio-cognitive achievement motivations, each of the six articles in this special issue focuses on a different motivational catalyst. These are values affirmation (Brady et al., 2016, pp. 353–373), growth mindset (Yeager et al., 2016, pp. 374–391), perceived similarity with others (Gehlbach et al., 2016, pp. 342–352), role models of persistence (Lin-Siegler, Ahn, Chen, Fang, & Luna-Lucero, 2016, pp. 314–328), performance and mastery goals (Park, Gunderson, Tsukayama, Levine, & Beilock, 2016, pp. 300–313), and strategies that help people avoid lapses in willpower (Duckworth, White, Matteucci, Shearer, & Gross, 2016, pp. 329–341).

Despite the complexity of motivation and the variety of possible interventions, there is an understudied, yet consistent, pattern across socio-cognitive interventions. The pattern is that the positive effects of the treatment are restricted to a subgroup of the students—mainly, those students marked at risk for poor academic achievement, whether based on prior achievement or broad demographic categories. The selective effect appears in many other studies as well (e.g., Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009; Walton & Cohen, 2007). Students marked as

Daniel L. Schwartz, Katherine M. Cheng, and Shima Salehi, Graduate School of Education, Stanford University; Carl Wieman, Graduate School of Education and Department of Physics, Stanford University.

Correspondence concerning this article should be addressed to Daniel L. Schwartz, Graduate School of Education, 485 Lasuen Mall, Stanford University, Stanford, CA 94305. E-mail: Daniel.schwartz@stanford.edu

tapraid5/zcz-edu/zcz-edu/zcz00316/zcz2768d16z | xppws | S=1 | 3/22/16 | 8:14 | Art: 2016-1329

most at risk are the ones who benefit from an intervention, whereas the unmarked students show almost no benefit.

A pattern across complexity can be a foothold into explanation, so it may be worth considering why these interventions, regardless of method and theory, regularly work for some students and not others. All the interventions are broad based—they deliver universal messages that are not tailored to a specific segment of the student population or a domain of knowledge. For example, consider interventions that help students learn to interpret failure as an opportunity for growth. As Lin-Siegler et al. (2016) state, "everyone fails sometimes." Everyone can relate to a message that you can grow through failure, yet the evidence indicates that only a select few benefit from the message, presumably those who experience failure a lot. There are also practical reasons for understanding why an intervention works for some people but not others. As Brady et al. (2016) state,

> A one-size-fits-all approach is apt to disappoint. Far more effective is allocating interventions to the people that will most likely benefit, as in medical science. A drug appropriately matched to a medical condition can save a life. But the same medication given to the wrong person, or for the wrong condition, could be iatrogenic. (p. 369)

In the warranted excitement over the positive benefits of socio-cognitive interventions, there have been limited efforts to understand the null effects for large subsets of students. Several authors have suggested that there should be future research on the question. For example, Gehlbach et al. (2016) propose, "Other studies could investigate whether the intervention might be adapted . . . to make it effective for all students rather than just a subset of students" (p. 350). Similarly, Lin-Siegler et al. (2016) state, "Future research should identify other individual differences among students that might also benefit from this intervention" (p. 323). Others have offered acknowledged ad hoc explanations for the selective effect. Yeager et al. (2016), for example, offer two distinct speculations for why high-achieving students do not benefit from their growth mindset treatment:

> Grades have a range restriction at the top, in part due to grade inflation. Furthermore, high-achieving students may not get better grades when they begin seeking out challenging problems that could give them a lower grade but that might teach them something. . . . (p. 387)

Brady et al. (2016) also describe two possibilities, in their case, for why White students did worse as a result of the values affirmation treatment.

> [N]egative effects may occur for non-stereotyped students because they remind them of valued domains other than academics, which could lead them to invest their efforts elsewhere. Perhaps, also, some subset of White students is performing well in school because of stress and psychological threat and alleviating this may be counterproductive. (p. 369)

Ultimately, it is inevitable that some socio-cognitive interventions will fail to demonstrate positive effects. Sometimes, it will be due to chance in human affairs. Other times, it will be due to subtle variations in implementation. In the meantime, there is already evidence that our current interventions do not work for large segments of the school population. Understanding why they do not work in these cases may provide some guidance for knowing where and when future interventions can be most effective.

Comments on the Four Socio-Cognitive Intervention Studies

Each article presents a major piece of empirical and theoretical work, and each deserves several pages of praise and scrutiny. However, given the confines of a commentary, we review the six articles in the special section with the following question in mind: Why do these broad-based interventions appear ineffective for a large proportion of students? In our analysis, we do not reach a definitive conclusion, though we offer some speculations at the end. At the same time, there are methodological considerations that merit attention in future studies. These include the quality of instrumentation and a subtle confirmation bias that directs attention away from discrepant results and those students who do not benefit from the interventions.

Table 1 provides a rough summary of the four intervention studies by separating the students into the two quasi-experimental groups described by the researchers. The table holds two results of interest. First, it shows the relative effect of the treatments on academic achievement. For the students to the left, the gain in grade point average (GPA) is always greater for treatment than control. For the students to the right, the GPA gain is never higher for treatment than control. Second, it shows the relative effect of the intervention on theoretically important mediating variables such as mindset. In a perfect and simple world, the relative direction of effect on these variables would parallel the direction of effect on GPA, which was not always the case. T1

The study by Brady et al. (2016), which applied a values affirmation intervention for students, makes a pair of important research contributions, beyond the demonstration of effectiveness for Latino/a students. It shows that values-affirmation interventions, in which students briefly write about their most important values, can be effectively administered outside of the classroom, which opens new possibilities for delivery. It also appended an interesting measure to the end of the study. The students listed their imminent academic due dates (a stressor) and then wrote about anything they wished. The authors coded the free write to determine whether students spontaneously introduced self-affirmations at a greater rate than they mentioned perceived threats. The written essay was a great addition to the standard arms-at-a-distance analysis of GPA gains, rates of attrition, and survey responses, because it provided a direct behavioral indicator of the psychological processing that the intervention put in place. Given the first-of-kind nature of the coding, it will be important to learn if the researchers' decisions for how to combine the different codes, such as subtracting perceived threats from self-affirmations, works in subsequent studies.

This article is perhaps the best demonstration of the mystery of selective effects. The Latino/a students showed a positive effect of the values affirmation treatment, which is an important replication. However, the White students, if anything, showed a negative effect of the treatment. This is a rare instance of a negative effect for "average" White students, so pending replication, we simply consider why the values affirmation did not help them, rather than why it may have hurt.

tapraid5/zcz-edu/zcz-edu/zcz00316/zcz2768d16z | xppws | S=1 | 3/22/16 | 8:14 | Art: 2016-1329

Table 1
A High-Level Summary of the Selective Effects of the Interventions

Study	Benefitting Population	Non-Benefitting Population
Brady et al. (2016)	Latino/a college students	White college students
Treatment:		
Students wrote about their most important value, and why it is important to them.	GPA gains: Treatment > Control	GPA gains: Treatment < Control
Control:		
Students wrote about 9th most important value, and why it is important for others.	Self-affirmations minus perceived threats: Treatment > Control	Self-affirmations minus perceived threats: Treatment = Control
Yeager et al. (2016)	Lowest achieving 9th graders	Higher achieving 9th graders
Treatment:		
Students read, reflected, and wrote about growth mindset.	GPA gains: Treatment > Control	GPA gains: Treatment = Control
Control:		
Students read, reflected, and wrote about transition to high school.	Extent of fixed mindset: Treatment = Control	Extent of fixed mindset: Treatment > Control
Gehlbach et al. (2016)	African American and Latino/a 9th graders	Asian American and White 9th graders
Treatment:		
Teacher learned about five commonalities with half their students.	GPA gains: Treatment > Control	GPA gains: Treatment = Control
Control:		
Teacher did not receive any information about other half of students.	#Teacher-Student interactions: Treatment > Control	#Teacher-Student interactions: Treatment = Control
Lin et al. (2016)	Lower achieving 9th graders	Higher achieving 9th graders
Treatment:		
Students read about famous scientists' intellectual or life struggles.	GPA gains: Treatment > Control	GPA gains: Treatment > Control
Control:		
Students read about famous scientists' achievements.	Extent of growth mindset: Treatment = Control	Extent of growth mindset: Treatment = Control

Note. GPA = grade point average.

Values affirmation is a very broad treatment. It is reasonable to have expected it to help White students who were feeling stressed about achieving the values of school. The White students who received the treatment had lower GPAs to start with than their control counterparts, so they presumably felt some academic threat coming into the study, and a values affirmation should have helped. Moreover, the White sample included women. Prior research has found that values affirmations can be beneficial for women in male stereotyped classes such as math (Miyake et al., 2010). Therefore, we would have expected to at least observe some improvements in the math scores of White women in the treatment condition.

One possibility is that there was an effect of the affirmation treatment on White students that was lost in the summary statistics of averages and slopes. It would have been helpful if the authors had separated out the GPA changes by gender for math courses specifically. Also, a smart graphic that represents the distribution of individual trajectories would be helpful for future reports. A second possibility is that whatever is going on with the psychology of the Latino/a students simply does not apply to the psychology of the White students, at least on average. The Latino/a students exhibited a stable relation between their tendency toward self-affirmation, their confidence, and their GPA gains. In contrast, the White students showed no treatment-relevant relations among these psychological processes that otherwise should have helped them handle academic stress and improve their GPA. It may be worthwhile to analyze the free-write transcripts of the White students separately to see if there are differences between the treatment and control students unique to the White students' experience of school.

In sum, this article is a useful contribution to the literature on self-affirmation, because it collected first-of-kind field data on the psychological processes that take hold after an affirmation intervention. The unusual demonstration of deleterious treatment effects on White students helps clarify that effects in social-behavioral sciences are rarely categorical and may only work some percentage of the time, just like most medical treatments only work some percentage of the time. A failure to replicate does not entail that a prior finding is spurious. More likely, it indicates that an intervention does not work every time, and on practical grounds, the field should try to determine the frequency a treatment is effective.

Yeager et al. (2016) conducted an ambitious growth mindset study implemented by a third party. The important contribution of this study is to show that socio-cognitive interventions are effective, even when not conducted by the researchers themselves. Students in treatment groups completed two 20-min computerized sessions on growth mindset over four weeks. Control groups received information about the transition to high school. After one semester, the intervention had helped the very lowest achieving students (the bottom fifth) improve their GPA, but showed minimal effect for others. The demonstration of effective hands-off scalability is a compelling argument for adopting the instructional package, especially for the lowest achieving students.

tapraid5/zcz-edu/zcz-edu/zcz00316/zcz2768d16z | xppws | S=1 | 3/22/16 | 8:14 | Art: 2016-1329

Why was the benefit of the treatment confined to the lowest students? It is unlikely there was a ceiling effect on grades. Students in the 60th percentile had room for improvement, and they did not improve. One natural hypothesis is that the higher achieving students already had a good mindset for school, so the intervention did not teach them anything that would have helped them further raise their grades. The data refute this hypothesis. The top 20% of students (GPA) had the largest improvement in their mindset and were more likely to choose hard problems. In contrast, there was no change in the bottom 20% of students in either mindset or challenge seeking behaviors.

The authors had several preregistered hypotheses that were not foregrounded in the written article, including one about changes to students' mindsets. "We expect that growth mindset composite scores (manipulation checks) will be significantly more changed in the treatment group as compared to the control group . . ." (retrieved from https://osf.io/aerpt/) The results did not fit this main effects prediction. The lowest achieving students did not exhibit changes to their mindsets and they did show achievement gains, whereas the opposite is true for the higher achieving students. By the rationale of preregistering hypotheses to create definitive a priori tests, the study "proved" the practical effects of the treatment for some students, and it "disproved" the psychological theory on which the intervention was based. This result magnifies that there remains considerable uncertainty as to exactly why these interventions work, suggesting a need for caution in carrying out large-scale implementations.

One possibility for the unexpected inverse relation between mindset gains and achievement gains is that the growth mindset lessons were taken up differently by the lower and higher achieving students. The higher achieving students, being good at school, learned what answers they should give to the mindset survey, but they treated the message of the intervention as just another thing to learn in school. In contrast, the lowest achieving students found a message of possible change compelling, and while they did not change their beliefs about intelligence, they did feel a boost of optimism that drove them forward. The authors may have data relevant to these hypotheses. The researchers had students complete a writing prompt after the second growth mindset lesson. It may be useful to examine student writing. According to our current speculation, the higher achieving students would reflect the instructional content literally, whereas the lower achieving students would exhibit hopefulness.

In sum, the strength of the article is an independently conducted study that showed a computerized growth mindset tutorial helps the lowest achieving students. The weakness, regarding our specific question, is that the authors maintained a strict methodology that prevented them from investigating why one of their hypotheses was falsified; many of the students improved their mindset but did not improve their achievement, and those who did improve their achievement did not change their mindset.

Gehlbach et al. (2016) conducted a clever within-class research intervention aligned with the contact hypothesis (Allport & Kramer, 1946); namely, getting people to know how they are the same will result in closer relationships and reduce conflict. In a 2×2 design, one factor was whether students learned about their personal similarities with their teacher. The other factor specified whether the teachers learned about their similarities with their students. The authors' technique for helping students and teachers get to know one another is a smart idea and easily scaled. It is important to note, however, that the research never actually tested the similarity hypothesis itself. Teachers either read about their similarity with some students or they read nothing at all. It may simply be that for teachers, learning something about their students—whether or not it is their shared similarities—is what is important. Now that the researchers have developed the basic methodology, they can begin to isolate the significance of similarity versus familiarity.

The authors found that for White and Asian American students, the intervention had no measurable effect. However, African American and Latino/a students did exhibit a benefit. In those cases, in which the teacher had learned about their similarities, African American and Latino/a students showed a trend toward improved grades relative to those students who were not revealed to the teachers. The authors were tentative about this finding, because it was post hoc. While we appreciate the risks of post hoc data exploration, the results were more convincing than those associated with their a priori hypotheses. Moreover, it is not obvious why readers should have more confidence in empirical results when authors claim a hypothesis beforehand, especially regarding a study that is not a replication of prior research. The authors' opinion before a study should not make the empirical results more or less convincing.

Why was the effect of the intervention confined to the African American and Latino/a students? One clue may be the demographic composition of the teachers in the study, 80% of whom were White. It is reasonable that White teachers already considered themselves similar to White students, so the treatment did not reveal anything the teachers did not already know. It would have been helpful to break out the similarity ratings by racial and ethnic categories to see if a sense of similarity increased more in the case of African American and Latino/a students, who presumably the White teachers took as less like themselves at the start.

The authors did collect a revealing self-report measure. They found that teachers who learned about their African American and Latino/a students reported that they subsequently interacted with them more. It is not a far leap to imagine that this increase in interaction explains the grade improvement. The teachers needed some help to initiate discussions with students from different backgrounds than themselves, in this case African American and Latino/a. For students whose backgrounds or academic values are similar to themselves, the teachers can assume common ground for initiating discussions; the learning of similarities would not help. The self-report data on teacher-student interaction also helps explain why letting teachers know about similarities made a difference, whereas letting students know of similarities did not. Teachers are in a better position to initiate discussions than students.

In sum, the strength of this article is the good idea for how to help students and teachers come to know one another. The work suggests that this is especially effective for those cases where students and teachers cannot assume commonality, which potentially diminishes their frequency of interaction. The weakness is that, while the authors did some exploratory analysis as to the explanation for why the White and Asian American students did not benefit, they could have delved into this question more deeply and presented relevant data.

tapraid5/zcz-edu/zcz-edu/zcz00316/zcz2768d16z | xppws | S=1 | 3/22/16 | 8:14 | Art: 2016-1329

Lin-Siegler et al. (2016) built on the basic insight that students may misinterpret their own struggles as signs of weakness. They took the interesting approach of using stories of famous scientists that are often included in science textbooks. They wrote some stories in the typical fashion to reflect the scientists' accomplishments. For example, they stated,

> In one textbook we reviewed, Albert Einstein was described as 'the most powerful mind of the twentieth century and one of the most powerful that ever lived . . . He was the most different from any other men . . .' (Hewitt, 2006, p. 715). (p. 323)

To parallel the stories about intellectual achievements, the authors also wrote stories that emphasized the scientists' intellectual or personal struggles. Students read one of these three types of narratives, and the authors compared students' academic beliefs, preferences, and science grades before reading the stories and several weeks later. The lowest achieving students—one standard deviation below the mean at pretest—showed an improvement in their grades if they read either of the two types of struggle stories compared to students who read the achievement stories. The high-achieving students, however, did not show an effect of treatment.

Instead of raw science grades, the authors relied on *z* scores, the average of which should be zero. If low *z* scores move upward to the mean, then other *z* scores need to come down. The comparisons between the conditions do not reveal absolute gains or losses as a result of treatment. Instead, they reflect the relative ranking of students from the three treatments. The implication is that we do not know if the struggle stories helped student performance or the achievement stories hurt student performance.

The authors implemented a series of standard and nonstandard surveys to capture the students' motivational profiles. Interestingly, students' beliefs about intelligence were not correlated with grades at any time. This points to the delicacy of these kinds of measures and the need for better theoretical understanding, given that in some studies these surveys do correlate with achievement measures and in other studies they do not. As the authors mention, their measures were insufficient to determine the mediating factors in this study.

The growth mindset measures also did not reveal a change from pre- to posttest. This finding is not too surprising. Students would need to make an inference from the portrayal of struggle to the idea of adopting a growth mindset. Getting students to draw inferences while reading is a challenging industry of educational research in its own right (e.g., getting students to self-explain; Chi et al., 1994). One simple explanation for the selective grade improvement is that the low achievers decided to try a little harder on the next few assignments, because working through difficulties had been vicariously modeled by the high status individuals in the stories.

Another possible explanation for the selective effect is that the high achieving students who read the struggle stories may not have identified with the struggles of the scientists, because they had not struggled recently, at least to any magnitude near the scientist stories. By this account, the high achieving students were just reading stories rather than tacitly connecting the messages of those stories to their own experience. Though the authors measured students' level of connection to the stories, we do not know how student's connectedness to the stories differed when broken out by low and high achieving students, so the presented data cannot address this speculation. It would have been nice if the authors had measured students' delayed recall of the stories to see what features of the stories were most salient to which students. Perhaps the memory of the high achieving students would be less filled with stories of struggle than the low achieving students.

In sum, the strength of this study is that the authors have opened up a goldmine of potential research into the function of scientist narratives within textbooks. The issue is no longer confined to representing a diversity of scientists; it now includes how these potential role models are described. The primary weakness of the study is that the measures of psychological change appear to be a mismatch for the likely effects of the intervention. We do not know if the higher achieving students failed to show an effect of treatment because, as the authors speculated, they struggled less and therefore the intervention was irrelevant, or whether they experienced plenty of struggle but they already had coping mechanisms for handling struggle, or possibly some other explanation entirely. The sole reliance on z-scores makes it particularly difficult to distinguish between different hypotheses.

Commentary on the Two Articles on the Etiology of Motivational Beliefs

All told, the four socio-cognitive interventions exhibited selective effects. The benefits of the interventions were confined to the very lowest achieving students in two studies and to select minorities in two studies. The next two studies provide evidence for the etiology of different attitudes and strategies relevant to motivation. Table 2 summarizes these studies. T2

Park et al. (2016) investigated changes over time in young students' academic beliefs, preferences, and abilities, and considered how these changes might relate to their teachers' theories of intelligence, content knowledge, and instructional practices. The study makes a very important finding. Teachers who overemphasize performance may have a negative influence on children's motivational profiles. However, whether or not a teacher emphasizes mastery does not appear to matter for young children's motivational profile. One can speculate that a performance orientation introduces an evaluative component to the children's school experience, which in turn, can trigger concerns about falling short and the possibility of negative reinforcement.

The authors created a motivational framework profile for first and second graders that combined two measures—attitudes toward intelligence (growth vs. incremental) and performance orientation (preference for easy or hard problems). The authors also used a standardized test to measure the children's applied math skills. These measures were taken twice, once in the fall at the start of the school year and once at the end of the year in the spring.

Students' motivational frameworks were correlated with their math skills at both the beginning and end of the year. Moreover, a higher motivational framework early on predicted greater gains in math skills; good motivation is a way to improve learning. Reciprocally, higher math skills early on predicted greater gains in motivational framework; success is a good way to improve self-efficacy. Worthy of note, the correlation between math skills at the start and end of the school year ($r = .8$) was far greater than any of the correlations between math skills and the student motivation variables ($rs < .4$). In the context of a special issue on motivations

tapraid5/zcz-edu/zcz-edu/zcz00316/zcz2768d16z | xppws | S=1 | 3/22/16 | 8:14 | Art: 2016-1329

Table 2
High-Level Summary of Two Articles on the Etiology of Motivational Beliefs

Study	Population
Park et al. (2016)	1st and 2nd-graders
Before having a given teacher	
Correlations among student math scores and beliefs	Math scores: High Positive Motivation > Low Positive Motivation
After having a given teacher	
Correlations among student math scores, beliefs, and teacher beliefs	Positive motivational frameworks: High Math > Low Math High performance-oriented < Low performance-oriented teacher instruction teacher instruction
Duckworth et al. (2016)	High school and college students
Treatment:	Goal accomplishment:
Students received situation modification training	
Control 1:	Treatment > Control 1
Students received response modulation training	Treatment > Control 2 Control 1 = Control 2
Control 2:	Temptation:
Students received no training	Treatment < Control 1 Treatment < Control 2 Control 1 = Control 2

that are often taken as topic-independent, this finding is a good reminder that topic knowledge still plays an extremely important role in students' future learning.

In the middle of the school year, the authors asked teachers to self-report their instructional practices (mastery and performance), theory of intelligence (growth vs. incremental), and pedagogical content knowledge in math. None of the measures correlated with each other. As mentioned, self-reported performance-oriented instructional practices correlated with negative changes in student motivational profiles (more entity frameworks with increased performance avoidance). Curiously, in a separate article that, like this study, also relied on female teachers, the authors found that female teachers' math anxiety was transmitted to female students such that female students were more affected than male students, suggesting a social modeling effect (Beilock et al., 2010). In the present study, we assume there were no gender specific effects, because the authors did not report any. If true, one might surmise that performance avoidance frameworks are not primarily communicated by social model, but rather, they take hold via the explicit judgmental signals communicated in performance-oriented instructional practice.

An interesting side note in the article is that teachers' performance-oriented instructional strategies were highly correlated across math and reading domains ($r = .90$), whereas mastery practices exhibited a much smaller correlation across domains ($r = .50$). The authors do not offer an explanation for this asymmetry. One possibility is that teachers' have a harder time executing mastery-oriented instruction in mathematics, because mathematics in early elementary school tends to have a clear right-and-wrong answer approach (though it does not have to). Thus, teachers tend to lean toward performance-oriented instruction, even if they believe in mastery pedagogies. If true, it highlights that instructional practices are not solely rooted in teacher's beliefs about intelligence and mastery. They are also found in the entrenched pedagogies used in certain domains, pedagogies that employ strong evaluative signals.

In sum, the strength of this study is the evidence that one way children develop poor motivational profiles is through their school environment, especially an environment that overemphasizes performance. The weakness, which is an acknowledged issue for the field, is that the measures of motivation are not nearly as well developed as measures of math skills. For instance, the authors combined children's intelligence frameworks and performance orientations into one measure, whereas they kept teachers' intelligence framework and performance-oriented instructional practices as separate measures.

Duckworth et al. (2016) focused on strategies that can help young adults overcome temporary lapses in motivation. The important contribution of this work is a compelling motivational framework based on the insight that people can change their environment to avoid moments when temptation is at its greatest and rationality is at its weakest. They showed that situational strategies are more effective than cognitive ones. An extreme instance of a situational strategy would be putting a lock on the refrigerator to prevent snacking. A cognitive strategy would be self-talk at the moment of temptation. In the context of articles that emphasize the effects of students' self-attributions (implicit self-talk), it is ironic that changing one's environment is the most effective way to achieve one's goals. Of course, it is hard to change one's school by putting a lock on the playground, for example. There is still a place for socio-cognitive interventions.

The authors found that students know and endorse situational strategies, but they do not spontaneously use them. Short training in situation modification strategies enhanced students' self-reported goal accomplishment. Students experienced less distracting temptation, which enabled them to focus on their goals. The research findings were self-report and correlational, but there has been experimental support using behavioral outcomes such as eating more fruits and vegetables (e.g., Oppezzo & Schwartz, 2013).

Given Duckworth and colleagues' (2016) findings, one possible explanation for why socio-cognitive interventions do not work for

tapraid5/zcz-edu/zcz-edu/zcz00316/zcz2768d16z | xppws | S=1 | 3/22/16 | 8:14 | Art: 2016-1329

many students is that the interventions do not teach any new strategies about how these students can achieve their goals more effectively. In a relevant small study, Schunk and Gunn (1985) taught division to 9- through 11-year-olds. Students could receive lessons to foster achievement motivation, lessons on good division strategies, both lessons, or neither lesson. Students who received both lessons showed the greatest gains in division skills and in self-efficacy. By itself, the motivation treatment did little to help students learn division. It also did not lead to strong self-efficacy gains, presumably because self-efficacy is undermined by a lack of skill to achieve desired outcomes. A useful line of research would determine whether complementing socio-cognitive motivational interventions with strategy tips would raise the full distribution of students and not only the students marked for academic challenges.

In sum, the strength of this article is that it introduces a new theory of motivational strategies that has great face validity. The weakness, for our question at least, is that it did not examine whether some students did not benefit from the training, and if not why.

Speculations on the Unexplained Selective Effects of Socio-Cognitive Interventions

Brief socio-cognitive interventions can have lasting positive effects on achievement, especially for students at risk for low achievement. This is a tremendously important and positive fact, and it points to the power of beliefs. Yet, many students do not appear to benefit from these interventions—a useful caveat for educational providers and parents of excelling students, who are considering product purchases.

Here, we speculate on one possible way to explain the lack of effect for some students. We assume that all students experience academic self-doubt and stress at some point, both those who benefit from the interventions and those who do not. A key difference is that some students are able to recover their momentum after stress without any intervention and return to their full capacity. A useful analogy is riding a bike. When just starting to pedal, it takes a great deal of energy and focus to get rolling; it is difficult to steer and look around; and, a small bump can knock one off track. In contrast, once somebody is rolling along, it is easier to maintain speed; it is easier to look around and steer; and, the bike self-corrects when hitting small bumps in the road. By this analogy, there are many students who are already rolling along. Regaining their self-integrity is a matter of course. They may have a better map of possible resources, solutions, and the time needed to recover. When they hit a bump, they can think about and execute positive courses of action rather than forward chaining through an imagined sequence of cascading failures. The opposite of poor achievement is not necessarily a growth mindset, a sense of belonging, a mastery attitude, or values that relieve the pressures of doing well. People may only need a subset of these or other motivational candidates to maintain sufficient momentum forward. In contrast, there are those students who do benefit. The treatments provide them with a way to reach a speed and trajectory that they previously did not have and that requires less energy to maintain, steer, and recover.

The purpose of this analogy is to explain the majority of students who do not benefit from the treatments and to clarify why the interventions may be ineffective for them. It is not our intent to imply that the students who benefit from the interventions are without extra material challenges, threats, or struggles. They may have to start pedaling uphill, or they may have more bumps in the road. Indeed, the articles by Lin et al. (2016), Park et al. (2016), and Gehlbach et al. (2016), point to the contribution of schools to the very problems they hope to solve. Duckworth et al. (2016) go even further in isolating the contribution of situations to goal-directed behavior, because they show how students can materially change it to their advantage.

There are far too many unknowns to move beyond analogy. Therefore, we end with some thoughts for future research. We start from the observation that many socio-cognitive intervention studies rely on theories and findings developed in the laboratory. The goal of the intervention research is often to see if the findings from the laboratory also hold in the real world. We have two thoughts about this research mindset.

First, when we consider field trials of a new drug, there are extensive provisions for collecting unanticipated side effects, not the least of which are regularly scheduled check-ins. It would be unconscionable to run a clinical trial by giving people a drug, losing contact for 1 year, and then coming back to only check the vital signs purportedly addressed by the drug. Given the frequent analogy made from medical trials to educational trials, it may be time for educational trials to take a closer look at what is happening to the participants. To do this, the field needs to meet its grand challenge—develop better instrumentation that can regularly capture relevant choices, behaviors, and attitudes over time, and how these vary across different academic domains and contexts. We should note that many of the authors have noted the limitations in their data and the opportunities for improvement, but in considering all the studies as a whole, it is clear that the need is global and not limited to any particular study or studies.

Second, when treatments move from the lab to the real world, people often conceptualize the question as follows: Can the treatment stand up to the uncontrolled variability of the real world? "Sure, getting threatened people to complete a self-affirmation improves their performance on an immediate test in the laboratory, but can it improve their learning in school?" When an experiment yields a positive real-world outcome, our instinct might be to view it as a successful field trial demonstrating the ecological validity of a laboratory-grown intervention. However, the real world is not merely uncontrolled variability. It includes causal forces that could never appear in a laboratory, yet may contribute to a treatment's in situ effectiveness. For instance, in the Brady et al. study, the Latino/a students benefitted from the values affirmation, whereas the White students did not. One possible real-world explanation is that the minority students could pick one another out in a crowd of the majority. They are a definable group. They talk to one another. They may talk to one another about their values—values that were recently reaffirmed. They create their own value supportive network. And so on. One can continue the story without our help. The point is that we have no idea whether this occurred. There is nothing in the current evidence one way or another.

We wonder if the push to conduct large intervention studies is potentially responsible for the lack of relevant data. Large randomized control trials include elaborate statistical methods and declarations to prevent spurious findings. But these do not prevent a confirmation bias from creeping into the research design, especially because they are trying to test for positive effects. This

tapraid5/zcz-edu/zcz-edu/zcz00316/zcz2768d16z | xppws | S=1 | 3/22/16 | 8:14 | Art: 2016-1329

confirmation bias is most evident in the lack of effort to understand why many students do not benefit from the interventions.

References

Allport, G. W., & Kramer, B. M. (1946). Some roots of prejudice. *The Journal of Psychology, 22,* 9–39. http://dx.doi.org/10.1080/00223980.1946.9917293

Beilock, S. L., & Gunderson, E., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences, 107,* 1860–1863. http://dx.doi.org/10.1073/pnas.0910967107

Brady, S. T., Reeves, S. L., Garcia, J., Purdie-Vaughns, V., Cook, J. E., Taborsky-Barba, S., . . . Cohen, G. L. (2016). The psychology of the affirmed learner: Spontaneous self-affirmation in the face of stress. *Journal of Educational Psychology, 108,* 353–373.

Chi, M. T., Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18,* 439–477.

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

Duckworth, A. L., White, R. E., Matteucci, A. J., Shearer, A., & Gross, J. J. (2016). A stitch in time: Strategic self-control in high school and college students. *Journal of Educational Psychology, 108,* 329–341.

Gehlbach, H., Brinkworth, M. E., King, A. M., Hsu, L. M., McIntyre, J., & Rogers, T. (2016). Creating birds of similar feathers: Leveraging similarity to improve teacher-student relationships and academic achievement. *Journal of Educational Psychology, 108,* 342–352.

Hewitt, P. G. (2006). *Conceptual physics*. San Francisco, CA: Pearson Addison Wesley.

Lin-Siegler, X., Ahn, J. N., Chen, J., Fang, F.-F. A., & Luna-Lucero, M. (2016). Even Einstein struggled: Effects of learning about great scientists' struggles on high school students' motivation to learn science. *Journal of Educational Psychology, 108,* 314–328.

Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science, 330,* 1234–1237. http://dx.doi.org/10.1126/science.1195996

Oppezzo, M., & Schwartz, D. L. (2013). A behavior change perspective on self-regulated learning with teachable agents. In R. Azevedo & V. Alevan (Eds.), *International handbook of metacognition and learning technologies* (pp. 485–500). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4419-5546-3_31

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

Schunk, D. H., & Gunn, T. P. (1985). Modeled importance of task strategies and achievement beliefs: Effect on self-efficacy and skill development. *The Journal of Early Adolescence, 5,* 247–258. http://dx.doi.org/10.1177/0272431685052008

Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology, 92,* 82–96. http://dx.doi.org/10.1037/0022-3514.92.1.82

Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., . . . Dweck, C. S. (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.

Received January 21, 2016
Revision received February 17, 2016
Accepted February 18, 2016 ■