Theoretical, Ethical, and Policy Considerations for Conducting Social–Psychological Interventions to Close Educational Achievement Gaps

Kevin R. Binning[*]
University of Pittsburgh

Alexander S. Browman[*]
Boston College

Social–psychological interventions in education have shown remarkable promise as brief, inexpensive, and powerful methods for improving educational equity and inclusion by helping underperforming students realize their potential. These findings have led to intensive study and replication attempts to understand and close achievement gaps at scale. In the present review, we identify several significant issues this work has raised that bear on the theoretical, ethical, and policy implications of using these interventions to close achievement gaps. Using both classic and contemporary models of threat and performance, we propose a Zone Model of Threat to predict when social–psychological interventions in education may yield positive, null, and negative effects for specific students. From this analysis, we argue from an ethical standpoint that to reduce backfire effects, interventions should be focused on optimizing the salience of psychological threat across students rather than on uniformly reducing it. As a long-term policy goal, intervention studies should follow a two-step process, in which students' individual levels of threat are first diagnosed and then interventions are tailored to the students based on their threat levels. Practical and theoretical implications of the proposed framework are discussed.

Theories of fixed group differences in intelligence have long been offered as explanations for group disparities in academic achievement (e.g., Hernstein &

---
[*]Correspondence concerning this article should be addressed to Kevin Binning, University of Pittsburgh, Pittsburgh, PA 15260. Tel: 412-624-9149 [e-mail: kbinning@pitt.edu] or Alexander Browman, Boston College, Chestnut Hill, MA 02467. Tel: 617-552-2761 [e-mail: browman@bc.edu].
Both authors contributed equally and first authorship is shared.

Murray, 2010; for historical reviews, see Fredrickson, 2015; Gould, 1996). For this reason, it is difficult to overstate the significance of studies demonstrating the decisive influence of context and transient psychological threats on academic performance (e.g., Brown & Day, 2006; Steele, 1988; Walton & Spencer, 2009). This work shows that, far from there being fixed and immutable differences in ability between racial-ethnic, gender, and social class groups, intellectual performance is malleable and dependent on social–psychological experience. When students are asked to prove their abilities on difficult, intellectual tasks (e.g., a standardized math test), group differences in performance tend to emerge in line with prevailing societal trends, with positively stereotyped students (e.g., Whites, men in math-centric fields) outperforming negatively stereotyped students (e.g., African Americans, women in math-centric fields). But when students' social–psychological realities are manipulated—for example, by telling them that the test was just a puzzle or otherwise not diagnostic of intelligence—the gaps in performance shrunk or disappeared altogether (for reviews, see Steele, 2011; Steele, Spencer, & Aronson, 2002). Lifting psychological barriers to performance allowed students to demonstrate their latent potential (Walton & Spencer, 2009).

Such findings presaged a variety of intervention research seeking to scale up and apply these lessons to address educational achievement gaps. Today, a growing number of large, longitudinal field experiments in education have documented social–psychological causes and tools to address group performance and achievement gaps in education. They have done so by changing students' subjective doubts and insecurities—changing the way students perceive and understand themselves, their peers, and their environment—and unlocking their potential in the process (for reviews, see Cohen & Sherman, 2014; Harackiewicz & Priniski, 2018; Walton & Wilson, 2018; Yeager & Walton, 2011). On the whole, the growing body of work shows incredible promise in translating mechanistic lab findings into brief, inexpensive, and powerful methods for improving not just test performance, but equity in educational outcomes more broadly. In one set of studies, African American seventh graders who completed a series of in-class self-affirmation writing exercises showed benefits on their academic performance that persisted 7–9 years later (e.g., greater college attendance; Goyer et al., 2017). Several other studies have shown that social–psychological interventions can bolster student outcomes (e.g., higher grade point averages [GPAs], fewer discipline incidents) for multiple years after the initial intervention (e.g., Binning, Wang, & Amemiya, 2018; Borman, Grigg, Rozek, Hanselman, & Dewey, 2018; Tibbetts et al., 2016; Yeager et al., 2016).

Such findings highlight a profound insight into the human condition: they show that seemingly small interventions can produce large, downstream consequences. As in the natural world, "Tall oaks from little acorns grow" (Johnson, 1841, p. 9; see Manke & Binning, 2015). But just as critically, the findings also reveal the power of intervention and, by extension, the power of those who implement

them. In the present review, we argue that *such power gives rise to our collective responsibility to understand both the breadth and boundaries of the effects.*

We believe that this responsibility arises for at least three reasons. The first is from the perspective of theory. As the science has proliferated, it has become clear that the effects of intervention are conditional, with positive results emerging in some studies, null results emerging in others, and even negative results emerging in some. Interventions are not "magic bullets" (Yeager & Walton, 2011) but instead operate in a specific theoretical range for participants who fit specific psychological criteria (e.g., those for whom negative academic stereotypes apply to their racial or gender group; Garcia & Cohen, 2013). However, there is a lack of clarity among the general research community about the requirements for successfully implementing various intervention approaches. This has led some in the media to question the value of psychological intervention at all (e.g., Denworth, 2019; Yong, 2016). From a theoretical perspective, then, understanding when and under what conditions various intervention techniques are likely to be effective versus ineffective, or even counterproductive, is critical for the research community's collective abilities both to evaluate intervention findings and to plan future research that builds systematically from the existing research.

Second, researchers have a responsibility to understand intervention effects from the perspective of policy. In recent years, there has been great enthusiasm on the part of parents, practitioners, policymakers, and the public for implementing low-cost psychological interventions in schools, classrooms, and the home (Haimovitz & Dweck, 2017; Yeager & Walton, 2011). Unfortunately, popular ideas from research are often scaled up and implemented "way ahead of how ready the science is" (D. S. Yeager, in Denworth, 2019), especially when a "fix" appears to be as simple as many social–psychological interventions appear. As a result, we believe that researchers also have a responsibility to be mindful and explicit about the policy implications and potential constraints on generality (Simons, Shoda, & Lindsay, 2017) of their intervention work, especially the caveats and unknowns regarding when a technique is likely to be effective versus ineffective if implemented in real-world academic settings.

Third, and related to the prior two reasons, is the ethical perspective: We argue that researchers have a responsibility to ensure that interventions in the field are being conducted in a way that minimizes potential harm and maximizes potential benefits for students and society more generally. As discussed, there is now very little doubt that interventions can have powerful benefits, but it is also clear that they do not always or inevitably do so (e.g., Chen, Usher, Brown, & Ford, 2019; de Jong, Jellesma, Koomen, & de Jong, 2016; Dee, 2014; Hanselman, Rozek, Grigg, & Borman, 2017; Protzko & Aronson, 2016). Treatment heterogeneity—the variation in the effects of a treatment across people or groups—is common in intervention research of all types (Mukerjee, 2015), and research in psychology is no exception (Baker, Skinner, & Redding, 2019; Hanselman et al., 2017; Tipton,

Yeager, Iachan, & Schneider, 2019; Van Bavel, Mende-Siedlecki, Brady, & Reinero, 2016).

Most notably, in the intervention literature, this heterogeneity includes cases where some students experiencing statistically significant negative or backfire effects on academic motivation and attainment (Binning et al., 2012; Canning, Priniski, & Harackiewicz, 2019; Durik & Harackiewicz, 2007; Durik, Hulleman, & Harackiewicz, 2014; Durik, Shechter, Noh, Rozek, & Harackiewicz, 2015; Priniski, Skiljevic, & Harackiewicz, 2019), as well as other approaches that have trended consistently negatively for some students across multiple studies (e.g., Goyer et al., 2017; Walton & Cohen, 2007). These findings are at least partly attributable to chance: If enough intervention studies are conducted, some are bound to yield negative results. However, here we present a model that specifies when such negative and background effects are likely to occur and, as such, the possibility of negative effects warrants careful consideration.

To summarize, in the present review, we argue that the existing literature addressing social–psychological interventions that target academic performance raises theoretical, ethical, and policy concerns that merit attention. We limit the present review to educational interventions that have an implicit or explicit goal of helping students overcome psychological barriers to high performance that emerge in the academic context. We review this evidence while presenting a theoretical framework that attempts to account for the potential for positive, null, and negative outcomes of these interventions. Our approach is descriptive of when and for whom different approaches may be most effective. We then conclude with several practical and policy recommendations for conducting social–psychological interventions targeting academic performance in the field ethically—that is, in a way that minimizes potential harm and maximizes potential benefits for students and society more generally.

A Zone Model of Threat for Predicting Academic Performance

We propose a *Zone Model of Threat (ZMT)* for understanding the relationship between psychological threat and academic performance. As displayed in Figure 1, our model is inspired by the Yerkes–Dodson relation between arousal and performance (Yerkes & Dodson, 1908), and it proposes that a student's level of performance in an academic setting is contingent on the level of psychological threat that they experience in that setting, such that the relationship between threat and performance follows a Gaussian function.

Drawing from three prominent theoretical frameworks that examine the antecedents of performance under challenging circumstances—social identity threat theory (Steele et al., 2002), the biopsychosocial model of challenge and threat (Blascovich & Mendes, 2000), and the expectancy-value theory of achievement motivation (Wigfield & Eccles, 2000)—we define psychological threat as a
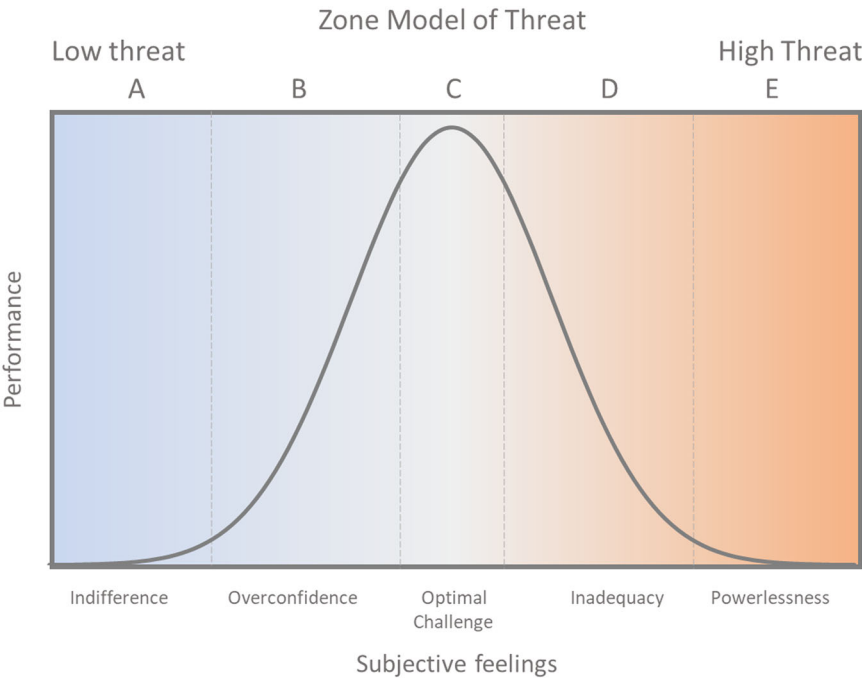
Fig. 1. Relationship between academic threat on academic performance theorized by the zone model of threat. The model assumes that threat is a product of two factors, the value or importance of the domain and students' expectancies to perform well in the domain. Optimal challenge (Zone C) occurs when value and expectancies are aligned. Feelings of inadequacy (Zone D) occur when value is high and expectancies are low, while feelings of overconfidence (Zone B) arise from low value and high expectancies. At extreme (Zones A and E) and optimal levels (Zone C), changes in threat may have null effects on performance.

product of two factors. The first is "how much the person identifies with" (Steele et al., 2002, p. 390) or *values the academic domain*: the more that "individuals view [academics] as central to their own sense of themselves" (Wigfield & Cambria, 2010, p. 39) and therefore that "one's self-regard, or some component of it, depends on [their academic] outcomes" (Steele et al., 2002, p. 390), the greater the sense of psychological threat that they will feel in an academic setting. The second factor is their *academic expectancies*, or "their beliefs about how well they will do [in school]" (Wigfield & Eccles, 2000, p. 68). Specifically, the less a student feels that they have sufficient personal or environmental resources to meet the demands of an academic situation, the greater the sense of psychological threat that they will feel in that situation (Blascovich & Mendes, 2000). Taken together, we therefore define psychological threat as a state of arousal that increases to the

extent that a student both identifies strongly with the academic domain and fears that they will not perform well therein (cf. Leary & Kowalski, 1997).

As a result, we propose that despite differences in the specific procedures and mechanisms they rely on, the multitude of existing social–psychological interventions designed to help improve students' academic performance can be viewed as addressing one of two broad categories of psychological barriers. The first are barriers that threaten students' *expectancies* about whether they can be successful in the academic domain. These include identity threats that emerge as a consequence of being the target of negative intellectual stereotypes (e.g., underrepresented minority students, female Science, Technology, Engineering, and Math [STEM] students, students from low-socioeconomic status [SES] backgrounds, first-generation university students; e.g., Johnson, Richeson, & Finkel, 2011; Mendoza-Denton, Downey, Purdie, Davis, & Pietrzak, 2002; Rheinschmidt & Mendoza-Denton, 2014; Walton & Cohen, 2007; see Steele et al., 2002), as well as nonidentity-based experiences that threaten a student's sense of academic competence (e.g., experiencing academic difficulty while believing that academic ability is innate and fixed; see Dweck, 1999; Walton, Paunesku, & Dweck, 2012).

We refer to approaches designed to address these barriers as *expectancy-enhancing interventions*, as they seek to bolster students' expectancies that their outcomes will comport with their motivations. These include (but are not limited to) self-affirmation interventions (Cohen, Garcia, Apfel, & Master, 2006; Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009; Harackiewicz et al., 2014), social belonging interventions (Walton & Cohen, 2007, 2011; Walton, Logel, Peach, Spencer, & Zanna, 2015), growth mindset interventions (Binning et al., 2019; Blackwell, Trzesniewski, & Dweck, 2007; Paunesku et al., 2015; Yeager et al., 2016, 2019), difference-education interventions (Stephens, Hamedani, & Destin, 2014; Stephens, Townsend, Hamedani, Destin, & Manzo, 2015; Townsend, Stephens, Smallets, & Hamedani, 2019), and attributional reframing interventions (Wilson & Linville, 1982, 1985). Notably, these interventions vary in how they enhance expectancies (e.g., by broadening self-views as in self-affirmation or by shaping meaning-making as in social belonging interventions), but we argue their common denominator is that they each seek to enhance expectancies.

The second category of interventions targets barriers that prevent students from seeing the academic domain as *valuable or personally important*. This can occur for many reasons. Students may see school or particular aspects of school as irrelevant to their lives and desired futures (see Oyserman & Destin, 2010). In other cases, students may experience overconfidence, for example, when they lack the expertise to appreciate their own academic limitations and knowledge gaps (e.g., Dunning, 2011; Dunning, Heath, & Suls, 2004). We refer to approaches designed to address these barriers as *value-enhancing interventions*. We propose these interventions are relevant to our threat-optimizing perspective because they function by increasing domain identification, or the extent to which "one's

self-regard . . . depends on the outcomes one experiences in the [academic] domain" (Steele et al., 2002, p. 390). This category includes (but is not limited to) utility-value interventions (Harackiewicz, Canning, Tibbetts, Priniski, & Hyde, 2016; Hulleman & Harackiewicz, 2009; Hulleman, Godes, Hendricks, & Harackiewicz, 2010) and self-transcendence interventions (Yeager, Henderson, et al., 2014), both of which aim to enhance domain identification by strengthening the psychological connection between academics and students' self-narratives and goals.

Together, this theoretical framework forms the basis of the ZMT's central prediction, illustrated in Figure 1: Performance on an academic task will be optimal when the student both identifies with the academic domain, and feels that they have adequate personal and environmental resources to feel challenged but able to meet the demands of the task (i.e., Zone C in Figure 1; see Blascovich & Mendes, 2000). Using optimal challenge as the reference point, threat becomes suboptimal when either expectancies are greater than value (e.g., when a student feels confident they can succeed but does not deem it important enough to try hard; Zones A and B) or when value is greater than expectancies (e.g., when students want to succeed in the domain but do not feel capable of doing so; Zones D and E). While the idea that there are optimal and suboptimal levels of threat for maximizing performance has been theorized and supported by the foundational frameworks discussed (Beilock, 2010; Blascovich & Mendes, 2000; Yerkes & Dodson, 1908), here, we apply this reasoning to understand the potential impact of social–psychological interventions to address educational achievement gaps.

The graphic representation in Figure 1 is mainly provided to clarify the general principles discussed here; we do not make strong claims about the precise shape of the function (e.g., in terms of kurtosis or whether the shape varies across people or situations). Rather, the core argument of the model is that the direction of the relationship between threat and performance varies across five theoretical zones of threat, such that threat and performance can be positively correlated (Zone B), relatively uncorrelated or weakly correlated (Zones A, C, and E), and negatively correlated (Zone D), depending on the starting or initial levels of threat. This nonlinearity occurs between people at given time points—with different groups having different mean levels of threat at a given moment in a given context (e.g., when African Americans feel greater threat than White Americans)—as well as within people over time—with value and expectancies varying within a given person over time (e.g., when students perform worse-than-expected in a domain and comes to devalue with the domain).

We therefore argue that interventions can be understood as efforts to move students to different points along the *x*-axis of Figure 1. But not all such movement is equal: A core implication of the model is that objectively identical increases or reductions in threat along the *x*-axis (e.g., −.5 *SD*) can result in very different implications for performance. Depending on which zone along the curve a student

starts, a reduction in threat (e.g., through expectancy-enhancing interventions) could lift performance for some students but leave it unchanged or even decrease performance for other students. We argue that the recognition of such underlying relationships generates several testable hypotheses for when interventions may produce positive, null, and negative effects.

Foremost, the *inadequacy zone* (Zone D) accounts for the prevailing mechanism believed to underlie many social–psychological interventions: They work when they undermine or transform psychological threats that would otherwise harm students' educational outcomes (e.g., Cohen & Sherman, 2014). This zone is consistent with the work of Steele (1997) in specifying that many interventions will be most impactful where academic threat is too high—that is, when a student is highly identified with the academic domain and sees them as personally important, but the situation leads the student to perceive academic success as unlikely (e.g., when stereotypes are "in the air"). In such cases, threat and performance are negatively correlated: when threat increases, performance decreases. Expectancy-enhancing interventions (e.g., via self-affirmation or social belonging; Cook et al., 2012; Walton & Cohen, 2007) boost performance by lifting expectancies to better optimize the relationship between expectancies and value, moving students toward Zone C and leading to higher predicted performance (see Figure 1).

However, the present model extends this perspective to understand how interventions can also produce null and negative effects. The simplest version of the model is that the familiar Goldilocks heuristic applies: some level of threat is beneficial, but too little or too much is not. This means that changes in threat levels—both reductions and increases in threat brought about by different interventions—can have varying consequences. It also means that interventions are unlikely to be most effective when they are implemented in a "one size fits all" manner. Different interventions may be needed based on the psychological states of participants at the time of intervention. In this review, we argue that expectancy-enhancing interventions are most useful when threat is too high (Zone D), whereas value-enhancing interventions are most useful when threat is too low (Zone B). We further argue that interventions may be ineffective when threat is extremely low (when value is extremely low; Zone A) and when it is extremely high (when expectancies are extremely low; Zone E). Furthermore, and perhaps less intuitively, we also posit that interventions may yield null effects among students for whom threat is already optimized (Zone C).

Unpacking the Five Zones of Threat

Zone C: Optimal Challenge

To begin, we consider what it means to experience optimal challenge, which represents Zone C of the ZMT. As discussed, we argue that optimal challenge

hinges on two factors: seeing *value* in the domain, and *expecting to succeed* in the domain. When these two conditions are met, students experience the feeling of having adequate resources to meet the valued academic work they face; they feel *optimally challenged* in academic settings.

The notion that there are optimal levels of threat for maximizing performance has been widely theorized and supported (Beilock, 2010; Blascovich & Mendes, 2000; Yerkes & Dodson, 1908): "moderate" levels of threat—when threat is neither too low nor too high—produce optimal performance. Furthermore, research on expectancy-value theory has shown that performance is highest when students both value (i.e., are highly identified with) academics and believe that they will be able to perform well in the academic domain (for reviews, see Wigfield & Eccles, 2000; Wigfield et al., 2015). For example, in two large-scale studies (Ns > 10,000) with pre-college students, Guo, Marsh, and colleagues (2015a, b) demonstrated that students who both (1) saw the most value in their academic work and (2) had the most confidence in their likelihood of being successful in school (corresponding to Zone C) performed better than both those who saw great value in academics but were less confident in their likelihood of success (i.e., Zone D) and those who were extremely confident in their likelihood of success but were less ego-identified with academics (i.e., Zone B).

Other studies also support the present thinking. In research that manipulated stereotype threat and test difficulty, O'Brien and Crandall (2003) showed that making stereotypes salient harmed minority students' performance, but only when the task difficulty was high. In other words, stereotypes were not always harmful for performance; they were only harmful when they created a contrast or perceived gap in students' ability to succeed. Furthermore, in work examining physiological stress responses (cortisol reactivity), Crum and colleagues found that participants who held a stress-is-enhancing mindset—that is, participants who were likely to make challenge rather than threat appraisals in response to stress—had more optimal physiological responses to stress (i.e., lowered cortisol reactivity among high reactivity participants and increased cortisol reactivity among low reactivity participants; Crum, Salovey, & Achor, 2013).

We argue that one of the defining features of being in the optimal challenge zone is that it makes performance less reactive to threat. For example, a .5 *SD* increase in threat for students in the optimal zone (e.g., from –.25 to .25 *SD* around the mean of 0) could leave predicted performance unchanged according to the ZMT. This implies that if one population of students is in a relatively optimal threat zone (e.g., historically privileged and over-represented groups), they should tend to show weaker intervention effects. Indeed, this tends to be the case when interventions have benefits in diverse settings. Majority group, continuing-generation, and male students typically exhibit small and often nonsignificant changes in performance following interventions, regardless of whether they were exposed to a self-affirmation intervention (Cohen et al., 2006, 2009; Harackiewicz

et al., 2014; Walton et al., 2015), a social-belonging intervention (Walton & Cohen, 2007, 2011; Walton et al., 2015), a utility-value intervention (Harackiewicz et al., 2016; Hulleman & Harackiewicz, 2009), a difference-education intervention (Stephens et al., 2014, 2015; Townsend et al., 2019), or a growth mindset intervention (Blackwell et al., 2007; Yeager, Johnson, et al., 2014; Yeager et al., 2019). However, when White students are under super-optimal levels of threat (e.g., from a low sense of belonging), they have been found to benefit from a self-affirmation intervention (Layous et al., 2017).

In addition, the notion that intervention effects will tend to be null among students already experiencing optimal levels of threat should be true not only on a group level, as evidenced above, but also within-people over time. Indeed, there is strong evidence for this in several intervention studies, particularly in work documenting how expectancy-enhancing interventions can change how people respond to threat over time. Research on "psychological tethering" among negatively stereotyped students provides one such example (e.g., Cook et al., 2012; Sherman et al., 2013; Walton & Cohen, 2007). Tethering refers to how two variables covary within a given person over time. For example, the salience of threat may vary within students over time, with feelings of threat being higher on some days (standardized test days) and lower on others (the day before a holiday break). Academic performance also varies within students over time, with students performing better-than-expected on some days and worse-than-expected on other days. If periods of high threat are associated with periods of lower performance, and days of low threat are associated with higher performance, then threat can be conceptualized as being "tethered" with performance within-persons, with the correlation between threat and performance being negative, as predicted for students experiencing feelings of inadequacy (Zone D). When students are in Zone C, these variables may be psychologically "untethered," such that changes in threat are uncorrelated with performance over time.

In one set of self-affirmation studies, for example, middle school students in the control condition showed a tethering effect: their variability in threat over time was negatively correlated with their GPA over time, such that periods of low subjective belonging were associated with periods of lower performance. This suggests that students in the control condition were in the inadequacy zone (Zone D) of the ZMT. By contrast, students hypothesized to be in Zone D (African American students) who received the affirmation intervention showed an untethering effect, such that for them, the variability in their experiences of belonging over time was unrelated to their academic performance over time (GPA; Cook et al., 2012). That is, consistent with the present argument, the intervention appeared to move students from Zone D, where threat and performance were negatively correlated, into Zone C, wherein changes in threat were uncorrelated with performance. This null relationship, we argue, is consistent with the flattening of the threat-performance curve within Zone C. These untethering results have also been

replicated in research involving self-affirmation and social belonging interventions (Sherman et al., 2013; Walton & Cohen, 2007).

Zone D: Inadequacy

Next, we contend that students experience feelings of inadequacy when their expectancies fall short of their value for the domain. It occurs when students want to do well but do not feel adequately equipped to do well. Stereotype threat is the paradigmatic threat for this zone (although but students from any background in a given context can experience feelings of inadequacy). Consistent with the present view of threat, stereotype threat occurs to the extent that students are (1) highly identified with and value academics but (2) feel that they do not have adequate resources to be successful in school (Steele et al., 2002). High levels of threat are cognitively taxing, making it difficult to concentrate and stay engaged on the very tasks that require concentration and engagement for optimal performance (Beilock, 2010; Schmader, Johns, & Forbes, 2008; Steele, 1997). As a result, for students in this zone, the relationship between threat and performance is negative: as threat increases, performance decreases. In the short term, stereotype threat taxes the cognitive resources students need for their academic tasks (see Schmader et al., 2008). In the long-term, underperformance may kick-start a downward cycle of recursive processes, such that lower performance raises concern about underperformance (i.e., expectancies), which further reduces performance (Cohen & Sherman, 2014).

As a result, we propose that interventions that lower perceived value and raise expectancies will diminish threat and increase performance for students in this zone. Among the first expectancy-enhancing interventions to gain attention in educational field settings were self-affirmation interventions, which were designed to reduce the effects of negative intellectual stereotypes on African American middle schoolers by having them write reflective essays about personally important values. In doing so, they were able to view threat from a broadened perspective—that is, being affirmed changed how threat was appraised (Cohen et al., 2006). This broadened perspective might decrease the perceived value of the importance of the domain and/or increase expectancies to perform in the domain.

Self-affirmation is just one example of an intervention that improved academic performance and closed achievement gaps by bolstering self-resources to contend with threat: a social belonging intervention can increase female engineering students' confidence in their ability to handle daily academic stressors (Walton et al., 2015); a growth mindset intervention may enhance students' beliefs that their intelligence can grow (Walton et al., 2012); a difference-education intervention can help first-generation college students see their typically stereotyped working-class backgrounds as a strength (Stephens et al., 2014, 2015); and an attributional framing intervention can help college freshmen see early academic setbacks at

college as normative and temporary (Wilson & Linville, 1982, 1985). Thus, these interventions work not by changing the sources of threat, but by bolstering the perceptions of adequacy to contend with the threats.

By the same logic, however, value-enhancing interventions may be expected to have the opposite effect for students who feel inadequate (Zone D), as making the domain feel even more important could increase threat to even less optimal levels. To illustrate, consider a student who sees academics as personally important, but they experience a feeling of inadequate resources to be successful in school. Meanwhile, a teacher or policymaker assumes underperformance occurs because students do not sufficiently value education (Puchner & Markowitz, 2015), which is a reasonable and important problem for some students but not all. The student therefore receives an intervention designed to increase the extent to which they see the academic domain as valuable and important. This would push the student further to the right on the *x*-axis of Figure 1, thereby further depressing their academic performance. Indeed, research has found that students with low perceived competence in math performed worse on math problems if they received information that promoted the utility value of math than if they did not (Canning & Harackiewicz, 2015; Durik et al., 2015). However, their performance improved if the utility value information was accompanied by a competence-boosting message (Durik et al., 2015), thereby bolstering their task expectancies. Together, the literature suggests that common expectancy-enhancing versus value-enhancing interventions may have opposing effects for students experiencing feelings of inadequacy (i.e., those in Zone D).

Zone B: Overconfidence

Steele and colleagues (2002) hinted at the distinction we draw between Zone D, which is marked by feelings of inadequacy, and Zone B, which is marked by feelings of overconfidence. As they explain, "It is likely that for the more strongly identified, the bigger part of their underperformance is mediated by the pressure of stereotype threat [and is therefore correctable by threat-reducing interventions] . . . But as . . . disengagement/disidentification become broader and more chronic, the pressure of stereotype threat may begin to play less of a performance-mediating role relative to that of low . . . motivation (p. 414)." Like Zone D, then, we propose that the subjective experience of being in Zone B is marked by an imbalance between values and expectancies. But rather than value outweighing expectancies, here expectancies outweigh value. Low value relative to expectancies results in overconfidence or undue optimism about one's adaptive adequacy in the academic domain.

This contention is noteworthy because research has shown that students who are sure that they can overcome any challenge in school are less likely to adequately prepare for those challenges (e.g., Nordgren, Van Harreveld, & Van Der Pligt,

Table 1. Theorized Effects on Performance for Expectancy-Enhancing and Value-Enhancing Interventions as a Function of Students' Zone of Threat at the Time of the Intervention

Examples of interventions	Student's zone prior to intervention				
	Zone A	Zone B	Zone C	Zone D	Zone E
Expectancy-enhancing					
Attributional reframing (Wilson & Linville, 1982)	∅	↓	∅	↑	∅
Difference-education (Stephens et al., 2014)	∅	↓	∅	↑	∅
Growth mindset (Blackwell et al., 2007)	∅	↓	∅	↑	∅
Self-affirmation (Cohen et al., 2006)	∅	↓	∅	↑	∅
Social-belonging (Walton & Cohen, 2007)	∅	↓	∅	↑	∅
Value-enhancing					
Self-transcendence (Yeager, Henderson, et al., 2014)	∅	↑	∅	↓	∅
Utility-value (Harackiewicz et al., 2016)	∅	↑	∅	↓	∅

Note: ↑ Performance boosting, ∅ Null performance effects, ↓ Performance sapping.

2009). Examples include students with a history of high performance who have not had to study or work very hard in school to earn good marks. One potential result is that these students may experience suboptimal levels of threat, and may therefore fail to take the necessary steps to prepare for or engage with academic tasks because they are overconfident in their abilities to succeed without additional preparation. Overconfident students (Zone B) are thus predicted to have a positive relationship between threat and performance, such that higher threat actually predicts higher performance.

As a result, interventions that enhance students' feelings of academic importance and value (e.g., utility value and self-transcendence interventions; see Table 1) should improve engagement and performance. To date, such interventions have often involved encouraging students to think about how what they are learning in school might "be useful to you, or a friend/relative, in daily life," "apply to your future plans," or "help you make the kind of impact you want on the people around you or society in general" (Harackiewicz et al., 2016; Hulleman & Harackiewicz, 2009; Hulleman et al., 2010). And critically, these approeaches have been shown to have positive, longitudinal effects both on the extent to which students see their schoolwork as interesting and important and on their GPAs.

By contrast, interventions that further reduce threat, moving students farther to the left on the *x*-axis, may actually harm performance. Consistent with this theorizing, Walton and Cohen (2007) found that White college students who were exposed to a social belonging intervention had lower-than-expected GPAs at the end of the academic year than did White students who were not exposed to the intervention. In other words, exposing historically high belonging students to an

intervention designed to enhance their confidence that they belong on campus had a negative effect on their academic performance, potentially by creating a sense of overconfidence or undue optimism about their adaptive adequacy at college. A similar finding was uncovered by Binning et al. (2018), who found that White adolescent students who reported high academic expectations at the beginning of the academic year actually had lower year-end GPAs if they were exposed to a mindset intervention than if they were assigned to a control condition consisting of mathematics practice.

Similarly, research has shown that self-affirmations can have disincentivizing effects, particularly for those who are not under high levels of psychological threat. Specifically, because affirmations highlight other domains in which students can invest their efforts, they may liberate students from feeling the need to behaviorally engage in activities of low interest or low perceived utility (e.g., engaging in repetitive or unchallenging tasks; Vohs, Park, & Schmeichel, 2013). Research by Kizilcec, Saltarelli, Reich, and Cohen (2017) provides a powerful example. This study examined behavioral persistence in massive online open courses (MOOCs) and compared students from more and less developed countries. For students from more developed countries, their experiences of threat and pressure were likely relatively low: widespread availability of the internet in their countries made the courses easy to access, and their higher levels of education meant that the stakes for successful completion of MOOCs were likely fairly low. For these students, then, exposure to a self-affirmation intervention may highlight that the course is not particularly essential or important, further reducing their experiences of threat and thereby freeing them to disengage from the course. Indeed, the authors found that for students from developed countries, exposure to a self-affirmation intervention actually *decreased* MOOC completion rates compared to those who were not exposed to the intervention.

Finally, although not a social–psychological intervention per se, efforts to indiscriminately enhance people's self-esteem are also instructive. Evidence from many studies indicates that increases in global self-esteem are tied to overall reductions in academic engagement and performance (Baumeister, Campbell, Krueger, & Vohs, 2003; Forsyth, Lawrence, Burnette, & Baumeister, 2007; Smith & Elliott, 2001; Zeigler-Hill et al., 2013). Thus, although potentially helpful to people with low self-esteem, across a population with moderate to high self-esteem, the overall impact of indiscriminate expectancy-enhancement seems to be non-positive.

Zone A: Indifference. Much like for students experiencing optimal threat (i.e., those in Zone C), we propose that if students are completely disengaged and therefore do not experience any amount of psychological threat in academic settings (i.e., students in Zone A), expectancy-enhancing and value-enhancing interventions are likely to be ineffective. This is because like those in Zone C,

students in Zone A are at one of the flattest points in the threat-performance curve, meaning that threat loses its predictive power in this zone. Although the value-enhancing interventions discussed in Zone B might be the best hope for these students (e.g., utility value and self-transcendence interventions), we contend that students in Zone A may be too indifferent or disengaged to engage with or even complete the intervention tasks. Indeed, a wealth of research has shown that a student must value and be highly identified with a domain—that is, the self must be "on the line"—in order to be motivated to work hard and perform well in that domain (see Steele, 1997; Steele et al., 2002; Wigfield & Eccles, 2000), and this should similarly apply to the interventions themselves.

Evidence of such null effects among disengaged students is seen in Durik, Harackiewicz, and their colleagues' work on utility value-enhancing frames (Durik & Harackiewicz, 2007; Shechter, Durik, Miyamoto, & Harackiewicz, 2011). Specifically, students were taught a new math technique under conditions where its utility value was either emphasized (e.g., how it could be used for doing personal banking or calculating tips at restaurants) or not. The manipulation boosted motivation among students who were more engaged with math (i.e., those high in math interest), but it was seemingly inert among students who were more disengaged (i.e., those low in math interest). Similarly, Borman and colleagues (2018) found that positive intervention effects from an expectancy-enhancing intervention only emerged among relatively engaged students.

Zone E: Powerlessness. Finally, interventions are also likely to be ineffective at very extreme levels of threat—that is, when students assign high value to the domain but have essentially no confidence in their ability to succeed (i.e., students in Zone E). In other words, we propose that Zone E is another flat part of the curve in which changes in threat lose their predictive power. Much like for indifferent students (Zone A), we propose that many interventions for students in Zone E will be too incongruous with their social–psychological situation to gain traction (see Walton & Yeager, in press). The key difference is that students in Zone E are highly motivated and engaged, but their lack of perceived or actual resources to be successful renders a feeling of powerlessness. This idea is similar to the concept of learned helplessness:

> [A] phenomenon in which repeated exposure to uncontrollable stressors results in individuals failing to use any control options that may later become available. Essentially, individuals are said to learn that they lack behavioral control over environmental events, which, in turn, undermines the motivation to make changes or attempt to alter situations (American Psychological Association, 2018).

In other words, students who are chronically exposed to uncontrollable stressors (e.g., lack of day-to-day life stability, threats to physical safety) or barriers (e.g., lack of environmental resources to support learning, chronic identity threats) in their academic environments may come to believe that they are academically

powerless—that they lack any control over their academic outcomes—making it unlikely that they would be motivated to try to make use of any "control options that may later become available," such as intervention-derived resources. In support of these predictions, students who are exposed to academic environments characterized by chronic uncontrollable stressors (e.g., chronic under-resourcing, school closures, aircraft noise, overcrowding) have been found to show patterns of helplessness on academic tasks (e.g., giving up sooner), which ultimately predicts weakened academic performance in the long term (e.g., Conner & Cosner, 2014; Evans, 2006; Fincham, Hokoda, & Sanders, 1989).

An additional example consistent with our argument is found in work on growth mindset interventions, which are designed to teach students that intellectual growth is possible. Such interventions have been shown to boost performance by reducing underperforming students' concerns about making mistakes and appearing incompetent, thereby enhancing their academic expectancies (Blackwell et al., 2007; Claro, Paunesku, & Dweck, 2016; Paunesku et al., 2015; Yeager et al., 2016; for review, see Yeager & Dweck, 2012). However, in a nationally representative study of over 12,000 ninth-grade students, Yeager and colleagues (2019, p. 364) found that such interventions are ineffective when students' academic environments do not "align with the messages of the intervention". In other words, if messages in the environment generally oppose or contradict the messages of a growth mindset intervention—such as when peers generally avoid academic challenges (Yeager et al., 2019), when teachers believe that intellectual ability is fixed (Canning, Muenks, Green, & Murphy, 2019), or when the person leading the intervention is deemed untrustworthy (Vaught & Castagno, 2008; Yeager, Purdie Vaughns, et al., 2014)—the intervention's benefits may be nullified (see Walton & Yeager, in press).

To summarize, interventions are likely to be ineffective in contexts where the academic climate makes students feel powerless to succeed. In fact, without concrete changes to the academic environment that could alleviate this sense of powerlessness, highly identified students experiencing such chronic and seemingly permanent environmental threats are likely to begin to engage in ego-defensive activities (or "acute reactions," as suggested by Steele et al., 2002)—such as domain avoidance (keeping away from a domain with which one is still ego-involved) and self-handicapping (preempting the negative evaluative consequences of performing poorly in an ego-involved domain)—and ultimately may even disengage or stop associating their self-regard with their academic outcomes altogether (a "chronic adaptation," as suggested by Steele et al., 2002). For example, when women who were committed to STEM were exposed either to an unbalanced (versus balanced) ratio of men to women attending a professional STEM conference or to less female STEM experts, they reported less desire to participate in the conference, identified less with STEM, and put in less effort on STEM tests (Murphy, Steele, & Gross, 2007; Stout, Dasgupta, Hunsinger, & McManus, 2011).

Similarly, African American college students who expected to take a diagnostic (versus non-diagnostic) test reported that they had slept less the night before, thereby preempting the negative evaluative consequences of performing poorly (Steele & Aronson, 1995).

In this way, indifference (Zone A) may be a long-term destination for students experiencing chronic powerlessness (Zone E). When high threat is experienced chronically, short-term disengagement efforts can result in a more permanent disidentification, or disconnection of one's self-regard from performance in the academic domain (Major, Spencer, Schmader, Wolfe, & Crocker, 1998; Wood-cock, Hernandez, Estrada, & Schultz, 2012). Because interventions in both of the extreme zones are likely to be ineffective, other remedies are needed to help these students. Along these lines, Yeager and colleagues (2019) call for "a new era of experimental research that seeks to enhance both students' mindsets *and the school environments that support student learning*" (p. 368, emphasis added; see also Walton & Yeager, in press).

The Potential for Backfire Effects: Zones B, C, and D

A key prediction of the ZMT is that interventions can produce backfire effects when they are delivered to students in the wrong zones of threat. Specifically, reducing threat among the disengaged or overconfident might make them even less engaged, while increasing threat among students who already feel inadequate may only make them feel more threatened. Both cases follow our argument that interventions should produce positive effects when they move students toward optimal levels of threat, but they should produce negative effects when they move students away from optimal levels of threat.

As a result, one circumstance in which interventions will produce backfire effects is when researchers or practitioners incorrectly anticipate how participants will interact and respond to the intervention stimuli. In self-affirmation interventions, for example, students are asked to reflect on personally important values, and research has shown that participants most frequently choose to affirm communal values like "spending time with family or friends" (Crocker, Niiya, & Mischkowski, 2008; Shnabel, Purdie-Vaughns, Cook, Garcia, & Cohen, 2013). However, research has also shown that for some students (e.g., first-generation college students), affirming these communal values can exacerbate feelings of conflict between their communal backgrounds and the individualistic university context, which ultimately impairs their performance (Tibbetts et al., 2016).

Similarly, interventions can sometimes ironically accomplish the opposite of its intended goal. As discussed, in research using utility value interventions, students are given the opportunity to find personal meaning and relevance in their school subject-matter. For example, a student might see the relevance of particular biological processes that relate to a family member's heart condition or their own

digestive or reproductive health (Harackiewicz et al., 2014, 2016). However, some students may have difficulty making these personal connections, and thus, rather than enhancing students' feelings of relevance to their life, they might conclude instead that it is not relevant to their life and actually perform worse (Canning et al., 2019; Priniski et al., 2019).

Finally, sometimes an intervention message itself may be problematic. For example, the initiation of pro-multiculturalism and diversity policies—policies that are designed to improve environmental expectancies among minoritized individuals—can increase feelings of threat among majority group members (Craig, Rucker, & Richeson, 2018; Dover, Major, & Kaiser, 2016; Wilkins, Hirsch, Kaiser, & Inkles, 2017) and provide a false sense of egalitarianism in the workplace (Dover, Kaiser, & Major, 2019). In the academic domain, for example, Browman and Destin (2016) found that while low-SES university students felt more comfortable on campus and reported greater athey reported seemingly-cademic expectancies when the university was seen as explicitly committed to supporting socioeconomic diversity, there were nonsignificant decreases for high-SES students, who typically report strong expectancies and feelings of belonging on campus. In other words, for students who rarely have to question their fit and belonging on campus, the message that students who are different from them also belong on campus may not only provide benefits to historically underrepresented students (i.e., moving them closer to the peak of the curve), but may also inadvertently decrease historically advantaged students' academic expectancies and thereby move them away from the peak of the curve.

Practical and Policy Implications of the ZMT

We propose that the ZMT provides three important practical contributions. First, it provides a theoretical framework for understanding how threat may contribute to performance differently for different students, and thus for understanding how the same intervention can be experienced differently by different students. Second, it suggests measurement strategies for pinpointing students' positions on the threat–performance curve. Finally, it offers a practical framework for helping practitioners tailor interventions to their students' threat zones.

Using the ZMT to Predict How Interventions Will Affect Different Students

A great majority of small-scale social–psychological interventions have been designed with an explicit goal of narrowing achievement gaps *by reducing threat among students who are assumed to be experiencing adaptive inadequacy as a result of threat* (i.e., moving them from in Zone D into Zone C; e.g., Cohen et al., 2006, 2009; Stephens et al., 2014; Walton & Cohen, 2007, 2011). However, as the field moves toward delivering interventions at scale—toward

testing "whether psychological interventions could practically be deployed to raise academic achievement. . . [for] virtually unlimited numbers of students at low marginal cost" (Paunesku et al., 2015, pp. 784, 790; see, e.g., Borman et al., 2018; Hanselman et al., 2017; Townsend et al., 2019; Yeager et al., 2016)—one issue the present model raises is that applying these threat-reducing interventions indiscriminately is not necessarily the best strategy to employ across an entire classroom, school, or district.

Specifically, small-scale threat reduction intervention studies are typically careful to disproportionately recruit and highlight the effects of the intervention on individuals who are assumed to be psychologically burdened by threat (i.e., students in the inadequacy zone [Zone D]; e.g., female and racial-ethnic minority students in stereotype threatening situations [Cohen et al., 2006; Walton et al., 2015], first-generation college students in culturally mismatched situations [Stephens et al., 2014]). By contrast, in the broader samples that scaling efforts seek to target, students will likely lie at various points on the threat curve, and therefore a given intervention will be unlikely to optimize the level of threat for all students.

Furthermore, even within groups with higher than average threat levels, many students may have initial mental states for which expectancy-enhancing interventions are ineffectual (e.g., feeling disengaged [Zone A], optimally challenged [Zone C], or powerless [Zone E]) or even harmful (i.e., feeling overconfident [Zone B]). For example, work by Jack (2014, 2016) highlights the striking heterogeneity of experiences that low-income Black students can experience on the same university campus. On one hand, several low-income Black university students attended under-resourced high schools in more distressed communities and reported experiences of academic threat and belonging uncertainty that suggest that they hold feelings of inadequacy at college, placing them in Zone D. One student noted, "[This college is] the first time I have ever been to school with Whites. My English class, I was the only black. My dance class, freshman year, I was the only black. The professor kept calling on me. I didn't feel comfortable" (Jack, 2014, p. 466).

By contrast, Jack also found that a large proportion of low-income Black students at the university (27%-37%) had high school experiences that were culturally and socially similar to their lives at university. Specifically, these students had attended highly resourced boarding, day, and preparatory schools in predominantly White and wealthy neighborhoods. Their university lives therefore greatly resembled their high school lives, and as a result, they reported seemingly more optimized levels of academic threat and belonging uncertainty on campus (Zone C). One student noted, "I fit in [at college because I was] a boarding school kid, getting some of the finest education. . . Sometimes I don't even think of myself as a low-income student" (Jack, 2014, p. 466). And others, still, reported characteristically extreme levels of threat that suggested powerlessness (Zone E)

or the beginnings of indifference (Zone A)—for example, "I closed myself off . . . sometimes [to] my professors. I was doing everything on my own . . . It made me miserable. Separating yourself from the environment while you're immersed in it, it causes internal conflict" (Jack, 2014, p. 468). In other words, "even for undergraduates from similar class [and racial-ethnic] backgrounds, the effects of . . . feeling like an outsider because of one's class [and racial-ethnic] background are not uniform" (Jack, 2014, p. 454). Large-scale applications of many existing interventions may not account for these students, despite ego-defensiveness and disengagement being a common outcome of chronic threat to one's personal and social status (Binning & Huo, 2012; (Major et al., 1998; Murphy et al., 2007; Steele & Aronson, 1995; Stout et al., 2011, Woodcock et al., 2012).

Adding further complexity, threat is often variable and unstable. Indeed, the salience of racial, gender, and SES-based prejudice, corresponding negative stereotypes, and feelings of identity safety may grow and recede depending on whether people think versus communicate about the prejudice (Binning & Sherman, 2011), the demographic composition of the social context (Binning & Unzueta, 2013; Borman, Grigg, & Hanselman, 2016; Browman & Destin, 2016; Murphy et al., 2007), the emergence or dissipation of credible messages of inclusion and acceptance (Purdie-Vaughns, Steele, Davies, Ditlmann, & Crosby, 2008; Murphy, Kroeper, & Ozier, 2018; Wanless, 2016), and other regional, historical, and cultural factors (e.g., national policy changes; Tankard & Paluck, 2017). Periods and experiences of high threat are therefore likely to be variable, with the same stressors having different subjective meanings at different time points (e.g., Cook et al., 2012; Sherman et al., 2013). Notably, it is this within-person malleability and context-specificity of experiences that opens the door for interventions to impact threat in the first place.

As a result, it is imperative that interventions that involve administration to large groups of students be sensitive to the heterogeneity in threat that may exist within and between populations of students, rather than focusing solely on addressing the prototypical Zone D student (Tipton et al., 2019). This view is particularly clear through an intersectional lens. For example, while certain intervention approaches have shown positive effects of women's performance in STEM classes on average (Miyake et al., 2010; Walton et al., 2015), they have not prioritized testing potential subgroup effects. For example, Asian women in STEM can potentially contend with negative stereotypes related to their gender and positive stereotypes related to their ethnicity. As a result, if an intervention designed to improve the performance of women in STEM ultimately activates the negative stereotypes for Asian women in the sample, it could undermine their performance (Shih, Pittinsky, & Ambady, 1999).

Similarly, although majority of group members are typically assumed to be in an optimal zone of threat, research has shown that they, too, can experience high levels of threat (e.g., Layous et al., 2017) or disengagement in the academic

context (e.g., Hu & Kuh, 2002). Furthermore, multiple studies that predominantly examined college students from racial-ethnic majority and high-SES backgrounds have found that the majority of participants rate themselves above average in intelligence, suggesting overconfidence (e.g., McCrae, 1990; Swann, Pelham, & Krull, 1989). Thus, while interventions using a uniform, expectancy-enhancing strategy may be helpful for students experiencing feelings of inadequacy (Zone D), they may potentially be inert for optimally challenged and indifferent students (Zones A and C), and may even be harmful for overconfident students (Zone B).

Using the ZMT to Estimate Students' Threat Levels

Given psychologists' ethical commitment to beneficence, researchers must attempt to determine if and how interventions may have helpful as well as harmful effects. We therefore propose that rather than focusing solely on group member-ships as a direct indicator of threat, efforts should also be made to identify where individual students actually lie along the threat continuum. As a result, we sug-gest some measurement strategies for doing so. Specifically, as discussed, each threat zone is characterized by a specific profile of psychological experiences or behaviors that we believe are unique to that zone: Students in Zone A experience indifference; those in Zone B are characterized by overconfidence; Zones C and D, by feelings that academics are important and that their resources for facing academic challenges are either adequate (Zone C) or inadequate (Zone D); and Zone E, by feelings of powerlessness and acts of ego-defensiveness. Each of these experiences and behaviors has been intensely investigated in prior research, and thus numerous validated self-report measures exist for assessing them.[1]

One strategy is therefore to assess each of these five psychological profiles in research samples prior to intervention in order to approximate students' positions on the threat-performance curve. Table 2 provides examples of some potentially useful measures, based on how we have conceptualized each zone; however, *this is by no means an exhaustive list*. First, they include only explicit self-report measures, which have limitations. Students may lack introspective awareness about the sources of their anxiety (e.g., Wilson, 2004). This means that a survey-based approach relying on self-reported attitudes may not provide a complete picture. Implicit methods have also been used as a strategy to gauge threat (e.g., word completion tasks, as in Cohen et al., 2006). These efforts used in conjunction with data analytics focused on where grade gaps emerge (e.g., between first-generation and continuing-generation students; between White and Black students;

---

[1] Psychophysiological measures also exist for some of these factors (see, e.g., Blascovich & Mendes, 2000; Stephens et al., 2015), but we do not discuss them here because they are less likely to be accessible to practitioners than self-report measures.

Table 2. Examples of Self-Report Measures Relevant to the Five Profiles of Psychological Experiences and Behaviors Hypothesized by the Zone Model of Threat

Zone and characterization	Sample measure	Sample item
Indifference and insufficient identity-involvement (Zones A and B) versus sufficient identity-involvement (Zones C, D, and E; reverse-scored)	Disengagement subscale measure (Major et al., 1998)	"It usually doesn't matter to me one way or the other how I do in school."
	Disengagement processes measure (Schmader, Major, & Gramzow, 2001)	"How I do intellectually has little relation to who I really am."
	Student-framed version of the identity subscale of collective self-esteem scale (Luhtanen & Crocker, 1992)	"Overall, being a student has very little to do with how I feel about myself."
Overconfidence (Zone B)	Positive academic illusions measure (Fischer et al., 2007)	"Compared to the average student at your school, how intelligent are you?"
	Academically-framed intellectual overconfidence subscale of the comprehensive intellectual humility scale (Krumrei-Mancuso & Rouse, 2016)	"For the most part, other students at my school have more to learn from me than I have to learn from them."
Adaptive adequacy[A] (Zone C) versus inadequacy[I] (Zone D)	Growth mindset measure[A] (Dweck, 1999)	"No matter how much intelligence you have, you can always change it quite a bit."
	Belonging measure[A] (Sherman et al., 2013)	"I feel like I belong at [school name]."
	Confidence about academic performance measure[A] (Harackiewicz et al., 2014)	"I am confident that I will do well in school"
	Belonging uncertainty measure[I] (Walton & Cohen, 2007)	"Sometimes I feel that I belong at University X, and sometimes I feel that I don't belong at [school name]."
	Academic and social concerns measure[I] (Sherman et al., 2009)	"I worry that people at my school will think I'm unintelligent if I do poorly"
Powerlessness (Zone E)	Academically-framed version of learned helplessness scale (Quinless & Nelson, 1988)	"I feel that I have little control over my outcomes at school."
	Academically-framed version of the personal constraints measure (Lachman & Weaver, 1998)	"I often feel helpless in dealing with the problems I encounter in school."

between men and women in certain STEM disciplines) may provide a powerful means to estimate students' zone of threat.

In addition, once quantitative measures have been administered, the data can then be analyzed using person-centered statistical methods (e.g., latent class

analysis) to infer what zones the students in their sample lie in, and therefore what types of intervention should be administered. For example, Matthews (2014) administered six measures of academic identity to a sample of African American and Latino adolescent males and found five distinct academic identity profiles, with notable implications for intervention (for another example, see Robinson, Perez, Carmel, & Linnenbrink-Garcia, 2019). For example, Matthews noted the following about students fitting two distinct profiles:

> "[Students fitting the first profile] do not think school is valuable/important, feel incompetent, and are estranged in school, [but] they are not quite disidentified, as they still evaluate school success as a parameter for how they feel about themselves (in terms of pride, respect, or satisfaction). Thus, one could speculate that . . . appropriate scaffolding to improve feelings of school belongingness, or developing sound regulatory strategies, would help these students re-engage academically . . . [By contrast, the second profile] may be particularly unique to marginalized males, as this population has historically internalized divergent messages, such as understanding the importance of school on an abstract level, but perceiving their incapability to be academically successful on a pragmatic level . . . [Thus] for this profile, a targeted intervention aimed at promoting high self-efficacy, through challenging work with acute scaffolding, would likely be impactful for increasing educational . . . performance" (p. 151).

While, to the best of our knowledge, this kind of profiling approach has not yet been used to tailor specific interventions to specific students, we believe that zone-relevant measures could provide a valuable basis for building profiles that could then be readily matched to specific threat optimizing interventions.

Using the ZMT to Select Appropriate Interventions

Finally, by identifying zones with unique threat-performance relationships, the model provides a framework for helping practitioners determine which interventions might be most beneficial in terms of moving students toward the peak of the curve. For example, both indifferent and overconfident students (Zones A and B) are characterized by lower levels of academic threat because they see little personal value in working hard in school. As a result, students in both zones might benefit from value-enhancing interventions designed to instill a realization that course material is relevant to their lives and is useful for achieving their personal goals (e.g., Harackiewicz, Hulleman, and colleagues' [2009, 2010, 2016] utility value intervention), or perhaps even goals that are bigger than themselves, like social justice (e.g., Yeager, Henderson, et al.'s [2014] transcendent self-purpose intervention). In the case of extremely disengaged students (Zone A), efforts must be made to engage the students in the intervention, such as by employing multiple value-promoting interventions or increasing their relevance to students via personalization (Browman, Svoboda, & Destin, in press; Canning & Harackiewicz, 2015; Rosenzweig et al., 2019).

By contrast, students experiencing feelings of inadequacy (Zone D) see academics as personally important but feel that they do not have the capabilities or resources to be successful. As a result, these students should benefit from one of the many expectancy-enhancing intervention approaches, depending on whether the barriers they experience stem from stereotype threat (e.g., self-affirmation interventions; Cohen et al., 2006, 2009; Harackiewicz et al., 2014, 2016; Walton et al., 2015), social belonging concerns (e.g., social belonging interventions; Walton & Cohen, 2007, 2011; Walton et al., 2015), cultural mismatch or environmental nonfit (e.g., difference-education interventions; Stephens et al., 2014, 2015; Townsend et al., 2019), or fixed mindsets (e.g., growth mindset interventions; Blackwell et al., 2007; Paunesku et al., 2015; Yeager et al., 2016, 2019). Relatedly, because powerless students (Zone E) still value academics but are so close to disengaging, it seems plausible that combined approaches that either target both students and external factors (e.g., contextual messages supporting inclusion [Murphy et al., 2018], teachers' mindsets [Canning et al., 2019]) or that incorporate both expectancy-enhancing and value-sustaining messages (Browman et al., in press) may lift barriers to threat reduction and improvements in performance.

Finally, for students in the optimal challenge zone (Zone C), social–psychological interventions may not be the best use of students' time. Academic skills training, such as training in metacognitive skills (Zepeda, Richey, Ronevich, & Nokes-Malach, 2015) or causal diagramming (Pearl, 2009), might be used instead to take advantage of the opportunity to challenge students who feel equipped and motivated to handle that challenge.

Tailoring Interventions to Particular Groups of Students

In support of our call for increased tailoring, research is increasingly showing the value of using data to deliver targeted interventions. For example, Dynarski, Libassi, Michelmore, and Owen (2018) conducted an intervention in which they partnered with the administration of a state's most selective university and sent personalized mailings to students encouraging to them to apply and promising 4 years of free tuition and fees if admitted. More specifically, they used state-level administrative data to target only rising seniors who were eligible for free- or reduced-price school meals and had sufficiently high test scores and grades for admission. In this way, the invention was able to increase application, enrollment, and attendance rates among low-SES students without affecting those of their higher-SES counterparts (see Dover et al., 2019).

As another example, a growing number of universities have begun offering mentoring programs for first-generation freshmen students (Plaskett, Bali, Nakkula, & Harris, 2018). Because these programs are designed to provide psychological resources specific to the needs of first-generation students, universities will typically target only those students for admission to the program. To

determine the effectiveness of the programs at narrowing the psychological gaps between first- and continuing-generation students, researchers can work with administrators to recruit samples of continuing-generation students who were not to be exposed to the program and have them complete survey measures and provide grades for comparison with students in the program. In one study utilizing this targeted approach, Browman, Destin, Cockrell, and Rivera (in preparation) found that one such program helped to narrow the gaps in perceived support and academic threat between first- and continuing-generation students.

It should be noted, however, that there are clear ethical concerns about visibly singling students out for interventions. If given only to certain students and under the guise of being "good for you," the intervention could itself be stigmatizing by insinuating to students that their differences are a weakness and therefore they are in need of help—a message that has been found to undermine the effectiveness of affirmations (Sherman et al., 2009; Silverman, Logel, & Cohen, 2013). By contrast, if an intervention instead emphasizes how their differences can be a source of strength (e.g., "I overcame the odds to be here, [and that] has prompted me to work harder and contribute more to [the university] now that I'm here"; Stephens et al., 2014, supplementary materials p. 6), this can promote better outcomes. Moreover, although there may be value in more personalized, tailored intervention strategies, recent ecological approaches—which attempt to intervene not on individual students but on the social context itself—have also shown benefits that extend to all students (Binning et al., in preparation; see also Walton & Yeager, in press).

Conclusions: Recommendations for the Future of Intervention Design and Testing

Taken together, the framework presented here—the ZMT—has several implications for the design and application of interventions in the field. Perhaps the most central recommendation is that neither expectancy- nor value-enhancing interventions should be uniformly administered to students without first establishing a thorough understanding of the threat levels of the students in the sample and the social context they inhabit. The occurrence of significant or trending adverse events noted in prior work suggests that in any heterogeneous treatment group, the same treatment may have varying effects on different students. This raises important ethical concerns regarding how we deliver interventions—specifically, that delivering interventions indiscriminately is not likely to yield the most optimal intervention effects (for related arguments, see Harackiewicz & Priniski, 2018; Rosenzweig et al., 2019; Yeager & Walton, 2011).

Our approach thus provides a number of useful recommendations for the future of intervention design and testing. First, a targeted approach hinges on developing a fuller understanding of how different interventions will influence

different students in different academic settings. As discussed, much of the research to date has focused on producing average effects with little consideration of the potential effect heterogeneity that an intervention could produce when applied to a diverse sample of students (see Tipton et al., 2019; but see, e.g., Borman et al., 2018; Yeager et al., 2019). One factor that we believe will be essential to overcoming this issue is an increase in the interdisciplinary nature of intervention research. For example, rigorous qualitative research by educational sociologists and psychologists has long been critical to advancing our understanding of the experiences of specific groups in academic settings (Jack, 2014, 2016; Lareau & Calarco, 2012; Matthews, 2018).

A powerful example of the value of interdisciplinary collaboration for the development of better interventions is the National Study of Learning Mindsets, conducted in the United States, which was designed to enhance understanding of "which kinds of students, in which kinds of classrooms, and in which kinds of schools are most likely to benefit from [growth mindset interventions]" (Mindset Scholars Network, 2019). This study involved collaboration among psychologists, sociologists, statisticians, economists, and education scholars to ensure that heterogeneity of growth mindset intervention effects could be properly tested. Consistent with the present argument, the results showed that growth mindset interventions do not benefit all students equally, and it helped isolate some necessary preconditions for tailoring (e.g., students are at increased risk of failure but their academic environment supports academic growth; Yeager et al., 2019).

Second, we advocate for an intervention strategy that first attempts to identify students' location on the threat-performance curve prior to intervention. By assessing the extent to which students are experiencing characteristics indicative of each zone (i.e., indifference, overconfidence, challenge, inadequacy, and power-lessness), we can work toward a better understanding of what interventions work when and for whom. In other branches of psychology, intervention research has developed to be adaptive, such that intervention treatments are delivered at only times when they are most needed (e.g., Nahum-Shani et al., 2019). Achieving such a level of precision in delivering social-psychological interventions is a worthy goal in light of the potential power of interventions.

Third, as discussed, future research must continue to work to identify important moderators of the effects of various interventions on students' position on the threat–performance curve (e.g., Borman et al., 2018; Ferrer & Cohen, 2018). For example, research has noted the importance of timing of interventions, which we interpret as the idea that the ability to move along the *x*-axis might be easier at particular times, such as immediately after a critical transition (e.g., to middle school or college) or before a difficult exam (Cohen & Sherman, 2014). In other words, threat may be more malleable at certain times and more fixed at others.

Finally, we recommend that future intervention efforts be more explicit about the constraints on the generality of their applicability (Simons et al., 2017). As discussed, not all members of a given social group will have had the same prior experiences (e.g., Jack, 2014, 2016), and the effects of a given intervention are dependent on these experiences (see Yeager & Walton, 2011). We have a responsibility to find out if and how broadly such hidden effects occur. There is currently a dearth of knowledge about the effects of certain interventions on students across the theorized zones. We therefore conclude by emphasizing that psychological interventions are only likely to be effective insofar as they address a psychological barrier that undermines performance. Knowing what contexts and situations impose which barriers for which students is a critical element of efforts to ethically and effectively deliver social–psychological interventions to foster equity in education.

References

American Psychological Association. (2018). *Learned helplessness*. Retrieved from https://dictionary. apa.org/learned-helplessness

Baker, D. J., Skinner, B. T., & Redding, C. H. (2019). Affirmative intervention to reduce stereo-type threat bias: Experimental evidence from a community college. *The Journal of Higher Education*. https://doi.org/10.1080/00221546.2019.1650582

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, *4*, 1–44. https://doi.org/10.1111/1529-1006.01431

Beilock, S. (2010). *Choke: What the secrets of the brain reveal about getting it right when you have to*. New York, NY: Simon and Schuster.

Binning, K. R., Cook, J. E., Purdie-Greenaway, V., Garcia, J., Chen, S., Apfel, N., . . . Cohen, G. L. (2019). Bolstering trust and reducing discipline incidents at a diverse middle school: How self-affirmation affects behavioral conduct during the transition to adolescence. *Journal of School Psychology*, *75*, 74–88.

Binning, K. R., & Huo, Y. J. (2012). Understanding status as a social resource. In K. Y. Törnblom & A. Kazemi (Eds.), *Handbook of social resource theory* (pp. 133–147). New York, NY: Springer.

Binning, K. R., Kaufmann, N., McGreevy, E., Fotuhi, O., Chen, S., Marshman, E., . . . Singh, C. (in preparation). Changing classroom contexts to foster equity and unlock student potential: An ecological belonging intervention. *Unpublished manuscript*.

Binning, K. R., & Sherman, D. K. (2011). Categorization and communication in the face of prejudice: When describing perceptions changes what is perceived. *Journal of Personality and Social Psychology*, *101*, 321–336.

Binning, K. R., & Unzueta, M. M. (2013). Perceiving ethnic diversity on campus: Group differences in attention to hierarchical representation. *Social-Psychological and Personality Science*, *4*, 500–507.

Binning, K. R., Wang, M.-T., & Amemiya, J. (2018). Persistence mindset among adolescents: Who benefits from the message that academic struggles are normal and temporary? *Journal of Youth and Adolescence*, *48*, 269–286. https://doi.org/10.1007/s10964-018-0933-3

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*, 246–263. https://doi.org/10.1111/j.1467-8624.2007.00995.x

Blascovich, J., & Mendes, W. B. (2000). Challenge and threat appraisals: The role of affective cues. In J. P. Forgas (Ed.), *Studies in emotion and social interaction, second series. Feeling and thinking: The role of affect in social cognition* (pp. 59–82). New York, NY: Cambridge University Press.

Borman, G. D., Grigg, J., & Hanselman, P. (2016). An effort to close achievement gaps at scale through self-affirmation. *Educational Evaluation and Policy Analysis*, *38*(1), 21–42.

Borman, G. D., Grigg, J., Rozek, C. S., Hanselman, P., & Dewey, N. A. (2018). Self-affirmation effects are produced by school context, student engagement with the intervention, and time: Lessons from a district-wide implementation. *Psychological Science*, *29*, 1773–1784. https://doi.org/10.1177/0956797618784016

Browman, A. S., & Destin, M. (2016). The effects of a warm or chilly climate toward socioeconomic diversity on academic motivation and self-concept. *Personality and Social Psychology Bulletin*, *42*, 172–187. https://doi.org/10.1177/0146167215619379

Browman, A. S., Destin, M., Cockrell, K., & Rivera, S. (in preparation). A mentoring program improves perceived support and reduces academic concerns for first-generation college students. *Manuscript in Preparation*.

Browman, A. S., Svoboda, R. C., & Destin, M. (in press). A belief in socioeconomic mobility promotes the development of academically motivating identities among low-socioeconomic status youth. *Self and Identity*. https://doi.org/10.1080/15298868.2019.1664624

Brown, R. P., & Day, E. A. (2006). The difference isn't black and white: Stereotype threat and the race gap on raven's advanced progressive matrices. *Journal of Applied Psychology*, *91*, 979–985.

Canning, E. A., & Harackiewicz, J. M. (2015). Teach it, don't preach it: The differential effects of directly-communicated and self-generated utility–value information. *Motivation Science*, *1*, 47–71. https://doi.org/10.1037/mot0000015

Canning, E. A., Muenks, K., Green, D. J., & Murphy, M. C. (2019). STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. *Science Advances*, *5*(2), eaau4734. https://doi.org/10.1126/sciadv.aau4734

Canning, E. A., Priniski, S. J., & Harackiewicz, J. M. (2019). Unintended consequences of framing a utility-value intervention in two-year colleges. *Learning and Instruction*, *62*, 37–48. https://doi.org/10.1016/j.learninstruc.2019.05.001

Chen, X.-Y., Usher, E. L., Brown, C. S., & Ford, C. (2019). *Brief social belonging intervention for first-year students at a land-grant university: Does modality matter?* In Annual Meeting of the American Education Research Association, Toronto, ON.

Claro, S., Paunesku, D., & Dweck, C. S. (2016). Growth mindset tempers the effects of poverty on academic achievement. *Proceedings of the National Academy of Sciences*, *113*, 8664–8668. https://doi.org/10.1073/pnas.1608207113

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*, 1307–1310. https://doi.org/10.1126/science.1128317

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. https://doi.org/10.1126/science.1170769

Cohen, G. L., & Sherman, D. K. (2014). The psychology of change: Self-affirmation and social-psychological intervention. *Annual Review of Psychology*, *65*, 333–371. https://doi.org/10.1146/annurev-psych-010213-115137

Conner, J., & Cosner, K. (2014). School closure as structural violence and stakeholder resistance as social justice. *Journal for Peace and Justice Studies*, *24*(2), 27–49. https://doi.org/10.5840/peacejustice201424216

Cook, J. E., Purdie-Vaughns, V., Garcia, J., & Cohen, G. L. (2012). Chronic threat and contingent belonging: Protective benefits of values affirmation on identity development. *Journal of Personality and Social Psychology*, *102*(3), 479–496. https://doi.org/10.1037/a0026312

Craig, M. A., Rucker, J. M., & Richeson, J. A. (2018). The pitfalls and promise of increasing racial diversity: Threat, contact, and race relations in the 21st century. *Current Directions in Psychological Science*, *27*, 188–193. https://doi.org/10.1177/0963721417727860

Crocker, J., Niiya, Y., & Mischkowski, D. (2008). Why does writing about important values reduce defensiveness? *Psychological Science*, *19*, 740–747. https://doi.org/10.1111/j.1467-9280.2008.02150.x

Crum, A. J., Salovey, P., & Achor, S. (2013). Rethinking stress: The role of mindsets in determining the stress response. *Journal of Personality and Social Psychology*, *104*, 716–733.

de Jong, E. M., Jellesma, F. C., Koomen, H. M. Y., & de Jong, P. F. (2016). A values-affirmation intervention does not benefit negatively stereotyped immigrant students in the Netherlands. *Frontiers in Psychology*, *7*, 1–20. https://doi.org/10.3389/fpsyg.2016.00691

Dee, T. S. (2014). Social identity and achievement gaps: Evidence from an affirmation intervention. *Journal of Research on Educational Effectiveness*, *8*(2), 149–168. https://doi.org/10.1080/19345747.2014.906009

Denworth, J. (2019, August). Debate arises over teaching "growth mindsets" to students. Scientific American. Retrieved from https://www.scientificamerican.com/article/debate-arises-over-teaching-growth-mindsets-to-motivate-students/

Dover, T. L., Kaiser, C. R., & Major, B. (2019). Mixed signals: The unintended effects of diversity initiatives. *Social Issues and Policy Review*. https://doi.org/10.1111/sipr.12059

Dover, T. L., Major, B., & Kaiser, C. R. (2016). Members of high-status groups are threatened by pro-diversity organizational messages. *Journal of Experimental Social Psychology*, *62*, 58–67. https://doi.org/10.1016/j.jesp.2015.10.006

Durik, A. M., & Harackiewicz, J. M. (2007). Different strokes for different folks: How individual interest moderates the effects of situational factors on task interest. *Journal of Educational Psychology*, *99*, 597–610. https://doi.org/10.1037/0022-0663.99.3.597

Durik, A. M., Hulleman, C. S., & Harackiewicz, J. M. (2014). One size fits some: Instructional enhancements to promote interest don't work the same for everyone. In K. A. Renninger & M. Nieswandt (Eds.), *Interest, the self, and K-16 mathematics and science learning* (pp. 49–62). Washington, DC: American Educational Research Association.

Durik, A. M., Shechter, O. G., Noh, M., Rozek, C. S., & Harackiewicz, J. M. (2015). What if I can't? Success expectancies moderate the effects of utility value information on situational interest and performance. *Motivation and Emotion*, *39*, 104–118. https://doi.org/10.1007/s11031-014-9419-0

Dunning, D. (2011). The Dunning-Kruger effect: On being ignorant of one's own ignorance. In J. Olson & M. P. Zanna, (Eds.), *Advances in experimental social psychology* (Vol. 44, pp. 247–296). New York, NY: Elsevier. https://doi.org/10.1016/B978-0-12-385522-0.00005-6

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, *5*(3), 69–106. https://doi.org/10.1111/j.1529-1006.2004.00018.x

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. New York, NY: Psychology Press.

Dynarski, S., Libassi, C. J., Michelmore, K., & Owen, S. (2018). *Closing the gap: The effect of a targeted, tuition-free promise on college choices of high-achieving, low-income students*. https://doi.org/10.3386/w25349

Evans, G. W. (2006). Child development and the physical environment. *Annual Review of Psychology*, *57*(1), 423–451. https://doi.org/10.1146/annurev.psych.57.102904.190057

Ferrer, R. A., & Cohen, G. L. (2018). Reconceptualizing self-affirmation with the trigger and channel framework: Lessons from the health domain. *Personality and Social Psychology Review*, *23*(3), 285–304.

Fincham, F. D., Hokoda, A., & Sanders, R. (1989). Learned helplessness, test anxiety, and academic achievement: A longitudinal analysis. *Child Development*, *60*, 138–145. https://doi.org/10.1111/j.1467-8624.1989.tb02703.x

Fischer, P., Greitemeyer, T., & Frey, D. (2007). Ego depletion and positive illusions: Does the construction of positivity require regulatory resources? *Personality and Social Psychology Bulletin*, *33*(9), 1306–1321. https://doi.org/10.1177/0146167207303025

Forsyth, D. R., Lawrence, N. K., Burnette, J. L., & Baumeister, R. F. (2007). Attempting to improve the academic performance of struggling college students by bolstering their self–esteem: An intervention that backfired. *Journal of Social and Clinical Psychology*, *26*, 447–459. https://doi.org/10.1521/jscp.2007.26.4.447

Haimovitz, K., & Dweck, C. S. (2017). The origins of children's growth and fixed mindsets: New research and a new proposal. *Child Development*, *88*, 1849–1859. https://doi.org/10.1111/cdev.12955

Fredrickson, G. M. (2015). *Racism: A short history*. Princeton, NJ: Princeton University Press.

Garcia, J., & Cohen, G. L. (2013). A Social Psychological Approach to Educational Intervention. In E. Shafir (Ed.), *Behavioral foundations of policy* (pp. 329–347). Princeton NJ: Princeton University Press.

Gould, S. J. (1996). *The mismeasure of man*. New York, NY: W. W. Norton.

Goyer, J. P., Garcia, J., Purdie-Vaughns, V., Binning, K. R., Cook, J. E., Reeves, S. J., . . . Cohen, G. L. (2017). Self-affirmation facilitates minority middle schoolers' progress along college trajectories. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(29), 7594–7599.

Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J. S., & Yeung, A. S. (2015a). Expectancy-value in mathematics, gender and socioeconomic background as predictors of achievement and aspirations: A multi-cohort study. *Learning and Individual Differences*, *37*, 161–168. https://doi.org/10.1016/j.lindif.2015.01.008

Guo, J., Parker, P. D., Marsh, H. W., & Morin, A. J. S. (2015b). Achievement, motivation, and educational choices: A longitudinal study of expectancy and value using a multiplicative perspective. *Developmental Psychology*, *51*(8), 1163–1176. https://doi.org/10.1037/a0039440

Hanselman, P., Rozek, C. S., Grigg, J., & Borman, G. D. (2017). New evidence on self-affirmation effects and theorized sources of heterogeneity from large-scale replications. *Journal of Educational Psychology*, *109*, 405–424. https://doi.org/10.1037/edu0000141

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Giffen, C. J., Blair, S. S., Rouse, D. I., & Hyde, J. S. (2014). Closing the social class achievement gap for first-generation students in undergraduate biology. *Journal of Educational Psychology*, *106*, 375–389. https://doi.org/10.1037/a0034679

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2016). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *Journal of Personality and Social Psychology*, *111*, 745–765. https://doi.org/10.1037/pspp0000075

Harackiewicz, J. M., & Priniski, S. J. (2018). Improving student outcomes in higher education: The science of targeted intervention. *Annual Review of Psychology*, *69*, 409–435. https://doi.org/10.1146/annurev-psych-122216-011725

Herrnstein, R. J., & Murray, C. (2010). *The bell curve: Intelligence and class structure in American life*. New York, NY: Simon and Schuster.

Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology*, *102*, 880–895. https://doi.org/10.1037/a0019506

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

Hu, S., & Kuh, G. D. (2002). Being (dis)engaged in educationally purposeful activities: The influences of student and institutional characteristics. *Research in Higher Education*, *43*(5), 555–575. https://doi.org/10.1023/A:1020114231387

Jack, A. A. (2014). Culture shock revisited: The social and cultural contingencies to class marginality. *Sociological Forum*, *29*, 453–475. https://doi.org/10.1111/socf.12092

Jack, A. A. (2016). (No) Harm in asking: Class, acquired cultural capital, and academic engagement at an elite university. *Sociology of Education*, *89*, 1–19. https://doi.org/10.1177/0038040715614913

Johnson, A. B. (1841). *The philosophical emperor: A political experiment*. New York, NY: Harper.

Johnson, S. E., Richeson, J. A., & Finkel, E. J. (2011). Middle class and marginal? Socioeconomic status, stigma, and self-regulation at an elite university. *Journal of Personality and Social Psychology*, *100*, 838–852. https://doi.org/10.1037/a0021956

Kizilcec, R. F., Saltarelli, A. J., Reich, J., & Cohen, G. L. (2017). Closing global achievement gaps in MOOCs. *Science*, *355*, 251–252. https://doi.org/10.1126/science.aag2063

Krumrei-Mancuso, E. J., & Rouse, S. V. (2016). The development and validation of the comprehensive intellectual humility scale. *Journal of Personality Assessment*, *98*(2), 209–221. https://doi.org/10.1080/00223891.2015.1068174

Lareau, A., & Calarco, J. M. (2012). Class, cultural capital, and institutions: The case of families and schools. In S. T. Fiske & H. R. Markus (Eds.), *Facing social class: How societal rank influences interaction* (pp. 61–86). New York, NY: Russell Sage Foundation.

Layous, K., Davis, E. M., Garcia, J., Purdie-Vaughns, V., Cook, J. E., & Cohen, G. L. (2017). Feeling left out, but affirmed: Protecting against the negative effects of low belonging in college. *Journal of Experimental Social Psychology*, *69*, 227–231.

Lachman, M., & Weaver, S. (1998). The sense of control as a moderator of social class differences in health and well-being. *Journal of Personality and Social Psychology*, *74*(3), 763–773.

Leary, M. R., & Kowalski, R. M. (1997). *Social anxiety*. New York, NY: Guilford Press.

Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, *18*(3), 302–318. https://doi.org/10.1177/0146167292183006

Major, B., Spencer, S., Schmader, T., Wolfe, C., & Crocker, J. (1998). Coping with negative stereotypes about intellectual performance: The role of psychological disengagement. *Personality and Social Psychology Bulletin*, *24*, 34–50.

Manke, K., & Binning, K. (2015, July). Mighty oaks from little (psychological) acorns grow. *SPSP Blog: Character & Context*. Retrieved from http://www.spspblog.org/mighty-oaks-from-little-psychological-acorns-grow/

Matthews, J. S. (2014). Multiple pathways to identification: Exploring the multidimensionality of academic identity formation in ethnic minority males. *Cultural Diversity and Ethnic Minority Psychology*, *20*, 143–155. https://doi.org/10.1037/a0034707

Matthews, J. S. (2018). When am I ever going to use this in the real world? Cognitive flexibility and urban adolescents' negotiation of the value of mathematics. *Journal of Educational Psychology*, *110*, 726–746. https://doi.org/10.1037/edu0000242

McCrae, R. R. (1990). Traits and trait names: How well is openness represented in natural languages? *European Journal of Personality*, *4*, 119–129.

Mendoza-Denton, R., Downey, G., Purdie, V. J., Davis, A., & Pietrzak, J. (2002). Sensitivity to status-based rejection: Implications for African American students' college experience. *Journal of Personality and Social Psychology*, *83*, 896–918. https://doi.org/10.1037/0022-3514.83.4.896

Mindset Scholars Network. (2019). *National study of learning mindsets*. Retrieved from https://mindsetscholarsnetwork.org/about-the-network/current-initatives/national-mindset-study/

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. https://doi.org/10.1126/science.1195996

Mukherjee, S. (2015). *The laws of medicine: Field notes from an uncertain science*. New York, NY: Simon and Schuster.

Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). Signaling threat. *Psychological Science*, *18*, 879–885. https://doi.org/10.1111/j.1467-9280.2007.01995.x

Murphy, M. C., Kroeper, K. M., & Ozier, E. M. (2018). Prejudiced places: How contexts shape inequality and how policy can change them. *Policy Insights from the Behavioral and Brain Sciences*, *5*(1), 66–74. https://doi.org/10.1177/2372732217748671

Nahum-Shani, I., Almirall, D., Yap, J. R. T., McKay, J. R., Lynch, K. G., Freiheit, E. A., & Dziak, J. J. (2019). SMART longitudinal analysis: A tutorial for using repeated outcome measures from SMART studies to compare adaptive interventions. *Psychological Methods*. https://doi.org/10.1037/met0000219

Nordgren, L. F., Van Harreveld, F., & Van Der Pligt, J. (2009). The restraint bias: How the illusion of self-restraint promotes impulsive behavior. *Psychological Science*, *20*, 1523–1528. https://doi.org/10.1111/j.1467-9280.2009.02468.x

Oyserman, D., & Destin, M. (2010). Identity-based motivation: Implications for intervention. *The Counseling Psychologist*, *38*, 1001–1043. https://doi.org/10.1177/0011000010374775

O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin*, *29*(6), 782–789. https://doi.org/10.1177/0146167203029006010

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

Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd edition). Cambridge, U.K. New York: Cambridge University Press.

Plaskett, S., Bali, D., Nakkula, M. J., & Harris, J. (2018). Peer mentoring to support first-generation low-income college students. *Phi Delta Kappan*, *99*, 47–51.

Priniski, S. J., Skiljevic, T., & Harackiewicz, J. M. (2019). *Designing utility-value interventions for first-generation college students*. In Annual Meeting of the American Education Research Association, Toronto, ON.

Protzko, J., & Aronson, J. (2016). Context moderates affirmation effects on the ethnic achievement gap. *Social-Psychological and Personality Science*, *7*, 500–507. https://doi.org/10.1177/1948550616646426

Puchner, L., & Markowitz, L. (2015). Do black families value education?: White teachers, institutional cultural narratives, & beliefs about African Americans. *Multicultural Education*, *23*, 9–16.

Purdie-Vaughns, V., Steele, C. M., Davies, P. G., Ditlmann, R., & Crosby, J. R. (2008). Social identity contingencies: How diversity cues signal threat or safety for African Americans in mainstream institutions. *Journal of Personality and Social Psychology*, *94*(4), 615–630.

Quinless, F. W., & Nelson, M. M. (1988). Development of a measure Of learned helplessness. *Nursing Research*, *37*(1), 11–15. https://doi.org/10.1097/00006199-198801000-00003

Rheinschmidt, M. L., & Mendoza-Denton, R. (2014). Social class and academic achievement in college: The interplay of rejection sensitivity and entity beliefs. *Journal of Personality and Social Psychology*, *107*, 101–121. https://doi.org/10.1037/a0036553

Robinson, K. A., Perez, T., Carmel, J. H., & Linnenbrink-Garcia, L. (2019). Science identity development trajectories in a gateway college chemistry course: Predictors and relations to achievement and STEM pursuit. *Contemporary Educational Psychology*, *56*, 180–192. https://doi.org/10.1016/j.cedpsych.2019.01.004

Rosenzweig, E. Q., Harackiewicz, J. M., Priniski, S. J., Hecht, C. A., Canning, E. A., Tibbetts, Y., & Hyde, J. S. (2019). Choose your own intervention: Using choice to enhance the effectiveness of a utility-value intervention. *Motivation Science*, *5*(3), 269–276. https://doi.org/10.1037/mot0000113

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*, 336–356. https://doi.org/10.1037/0033-295X.115.2.336

Schmader, T., Major, B., & Gramzow, R. H. (2001). Coping with ethnic stereotypes in the academic domain: Perceived injustice and psychological disengagement. *Journal of Social Issues*, *57*(1), 93–111. https://doi.org/10.1111/0022-4537.00203

Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science*, *12*, 1123–1128. https://doi.org/10.1177/1745691617708630

Shechter, O. G., Durik, A. M., Miyamoto, Y., & Harackiewicz, J. M. (2011). The role of utility value in achievement behavior: The importance of culture. *Personality and Social Psychology Bulletin*, *37*, 303–317. https://doi.org/10.1177/0146167210396380

Sherman, D. K., Cohen, G. L., Nelson, L. D., Nussbaum, A. D., Bunyan, D. P., & Garcia, J. (2009). Affirmed yet unaware: Exploring the role of awareness in the process of self-affirmation. *Journal of Personality and Social Psychology*, *97*, 745–764.

Sherman, D. K., Hartson, K. A., Binning, K. R., Purdie-Vaughns, V., Garcia, J., Taborsky-Barba, S., . . . Cohen, G. L. (2013). Deflecting the trajectory and changing the narrative: How self-affirmation affects academic performance and motivation under identity threat. *Journal of Personality and Social Psychology*, *104*, 591–618. https://doi.org/10.1037/a0031495

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, *10*, 80–83.

Shnabel, N., Purdie-Vaughns, V., Cook, J. E., Garcia, J., & Cohen, G. L. (2013). Demystifying values-affirmation interventions: Writing about social belonging is a key to buffering against identity threat. *Personality and Social Psychology Bulletin*, *39*, 663–676. https://doi.org/10.1177/0146167213480816

Silverman, A., Logel, C., & Cohen, G. L. (2013). Self-affirmation as a deliberate coping strategy: The moderating role of choice. *Journal of Experimental Social Psychology*, *49*, 93–98. https://doi.org/10.1016/j.jesp.2012.08.005

Smith, L. L., & Elliott, C. H. (2001). *Hollow kids: Recapturing the soul of a generation lost to the self-esteem myth*. New York, NY: Random House.

Steele, C. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181–227). San Diego, CA: Academic Press. https://doi.org/10.1016/S0065-2601(08)60229-4

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811. https://doi.org/10.1037/0022-3514.69.5.797

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*, 613–629. https://doi.org/10.1037/0003-066X.52.6.613

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, (Vol. 34, pp 379–440). NewYork: Academic Press. https://doi.org/10.1016/S0065-2601(02)80009-0

Steele, C. M. (2011). *Whistling Vivaldi: And other clues to how stereotypes affect us (issues of our time)*. New York, NY: WW Norton & Company.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and identity threat. *Advances in Experimental Social Psychology*, *34*, 379–440. https://doi.org/10.1016/S0065-2601(02)80009-0

Stephens, N. M., Hamedani, M. G., & Destin, M. (2014). Closing the social-class achievement gap: A difference-education intervention improves first-generation students' academic performance and all students' college transition. *Psychological Science*, *25*, 943–953. https://doi.org/10.1177/0956797613518349

Stephens, N. M., Townsend, S. S. M., Hamedani, M. G., Destin, M., & Manzo, V. (2015). A difference-education intervention equips first-generation college students to thrive in the face of stressful college situations. *Psychological Science*, *26*, 1556–1566. https://doi.org/10.1177/0956797615593501

Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. a. (2011). STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology*, *100*, 255–270. https://doi.org/10.1037/a0021385

Swann, W. B., Pelham, B. W., & Krull, D. S. (1989). Agreeable fancy or disagreeable truth? Reconciling self-enhancement and self-verification. *Journal of Personality and Social Psychology*, *57*, 782–791. https://doi.org/10.1037/0022-3514.57.5.782

Tankard, M. E., & Paluck, E. L. (2017). The effect of a supreme court decision regarding gay marriage on social norms and personal attitudes. *Psychological Science*, *28*, 1334–1344. https://doi.org/10.1177/0956797617709594

Tibbetts, Y., Harackiewicz, J. M., Canning, E. A., Boston, J. S., Priniski, S. J., & Hyde, J. S. (2016). Affirming independence: Exploring mechanisms underlying a values affirmation intervention for first-generation students. *Journal of Personality and Social Psychology*, *110*, 635–659. https://doi.org/10.1037/pspa0000049

Tipton, E., Yeager, D. S., Iachan, R., & Schneider, B. (2019). Designing probability Samples to study treatment effect heterogeneity. In P. J. Lavrakas (Ed.), *Experimental Methods in Survey research: Techniques that combine random sampling with random assignment* (pp. 435–456). New York, NY: Wiley.

Townsend, S. S. M., Stephens, N. M., Smallets, S., & Hamedani, M. G. (2019). Empowerment through difference: An online difference-education intervention closes the social class achievement gap. *Personality and Social Psychology Bulletin*, *45*, 1068–1083. https://doi.org/10.1177/0146167218804548

Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(23), 6454–6459.

Vaught, S. E., & Castagno, A. E. (2008). "I don't think I'm a racist": Critical race theory, teacher attitudes, and structural racism. *Race Ethnicity and Education*, *11*, 95–113. https://doi.org/10.1080/13613320802110217

Vohs, K. D., Park, J. K., & Schmeichel, B. J. (2013). Self-affirmation can enable goal disengagement. *Journal of Personality and Social Psychology*, *104*, 14–27. https://doi.org/10.1037/a0030478

Wanless, S. B. (2016). The role of psychological safety in human development. *Research in Human Development*, *13*, 6–14.

Walton, G. M., Paunesku, D., & Dweck, C. S. (2012). Expandable Selves. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 141–154). New York: Guilford Press.

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. https://doi.org/10.1037/0022-3514.92.1.82

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

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2015). Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering. *Journal of Educational Psychology*, *107*, 468–485. https://doi.org/10.1037/a0037461

Walton, G. M., & Spencer, S. J. (2009). Latent ability: Grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science*, *20*, 1132–1139.

Walton, G. M., & Wilson, T. D. (2018). Wise interventions: Psychological remedies for social and personal problems. *Psychological Review*, *125*, 617–655. https://doi.org/10.1037/rev0000115

Walton, G. M., & Yeager, D. S. (in press). Seed and Soil: Psychological Affordances in Contexts Help to Explain Where Wise Interventions Succeed or Fail. *Current Directions in Psychological Science*.

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

Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review*, *30*, 1–35. https://doi.org/10.1016/j.dr.2009.12.001

Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W., & Schiefele, U. (2015). Development of achievement motivation and engagement. In M. E. Lamb & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science* (7th ed., pp. 657–700). New York, NY: Wiley.

Wilson, T. D. (2004). *Strangers to ourselves*. Cambridge, MA: Harvard University Press.

Wilson, T. D., & Linville, P. W. (1982). Improving the academic performance of college freshmen: Attribution therapy revisited. *Journal of Personality and Social Psychology*, *42*, 367–376.

Wilson, T. D., & Linville, P. W. (1985). Improving the performance of college freshmen with attributional techniques. *Journal of Personality and Social Psychology*, *49*, 287–293.

Wilkins, C. L., Hirsch, A. A., Kaiser, C. R., & Inkles, M. P. (2017). The threat of racial progress and the self-protective nature of perceiving anti-White bias. *Group Processes & Intergroup Relations*, *20*, 801–812. https://doi.org/10.1177/1368430216631030

Woodcock, A., Hernandez, P. R., Estrada, M., & Schultz, P. W. (2012). The consequences of chronic stereotype threat: Domain disidentification and abandonment. *Journal of Personality and Social Psychology*, *103*(4), 635–646. https://doi.org/10.1037/a0029120

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, *47*, 302–314. https://doi.org/10.1080/00461520.2012.722805

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., ... Paunesku, D. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, *573*, 364–369.

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

Yeager, D. S., Johnson, R., Spitzer, B. J., Trzesniewski, K. H., Powers, J., & Dweck, C. S. (2014). The far-reaching effects of believing people can change: Implicit theories of personality shape stress,

health, and achievement during adolescence. *Journal of Personality and Social Psychology*, *106*, 867–884. https://doi.org/10.1037/a0036335

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. https://doi.org/10.1007/s13398-014-0173-7.2

Yeager, D. S., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., ... Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General*, *143*, 804–824. https://doi.org/10.1037/a0033906

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

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*, 459–482. https://doi.org/10.1002/cne.920180503

Yong, E. (2016, September). A worrying trend for psychology's 'simple little tricks.' *The Atlantic*. Retrieved from https://www.theatlantic.com/science/archive/2016/09/can-simple-tricks-mobilise-voters-and-help-students/499109/

Zeigler-Hill, V., Li, H., Masri, J., Smith, A., Vonk, J., Madson, M. B., & Zhang, Q. (2013). Self-esteem instability and academic outcomes in American and Chinese college students. *Journal of Research in Personality*, *47*, 455–463. https://doi.org/10.1016/j.jrp.2013.03.010

Zepeda, C. D., Richey, J. E., Ronevich, P., & Nokes-Malach, T. J. (2015). Direct instruction of metacognition benefits adolescent science learning, transfer, and motivation: An in vivo study. *Journal of Educational Psychology*, *107*, 954–970.