Trends in Cognitive Sciences

CellPress
REVIEWS

Spotlight

When Do Growth Mindset Interventions Work?

David I. Miller[1,*]

Can teaching students about brain plasticity improve their grades? A recent large, national experiment (Yeager *et al.*) found that a brief growth mindset intervention improved lower-achieving adolescents' grades by 0.10 points. Debate about interpreting the study's findings illustrates the need to consider effect heterogeneity and contextual factors when evaluating effect sizes.

The idea that brain functioning and intellectual abilities are malleable has captured the public's attention in recent years, leading many schools and districts to deliver 'growth mindset' interventions for fostering this belief among students [1]. Per mindset theory, students who believe their abilities can grow could embrace adaptive beliefs and behaviors (e.g., viewing failures as temporary setbacks, trying new study strategies, seeking help), improving their achievement over time, as some prior large-scale experiments have suggested [2,3].

Critics, however, have argued that the intervention effects are too small and unreliable to be practically meaningful, pointing to nonsignificant results in some replication attempts such as one recent, large UK experiment that trained teachers to deliver growth mindset lessons [4]. 'Resources might be better allocated elsewhere than mind-set interventions', argued the authors of a 2018 meta-analysis of 43 randomized controlled trials (RCTs) that evaluated intervention efficacy [1]. At the heart of this debate also lies deep theoretical questions about when, how, and how much motivational factors can influence academic and cognitive functioning in real-life settings [5,6].

A new large, preregistered national RCT (12 490 ninth-graders across 65 schools) aimed to inform these discussions by studying the conditions under which a carefully designed mindset intervention can improve grades [7]. Among its methodological strengths, the study raised the bar for evaluating the generalizability of educational intervention effects: it used stratified random sampling and survey weights to ensure national representativeness.

The brief mindset intervention (lasting less than 1 hour) was administered online with no prior teacher training and taught ninth-graders that the brain is like a muscle that gets stronger when used. Students also wrote about how they planned to apply the lessons to their own academic pursuits and study strategies. Based on individual-level random assignment, students received this intervention or an active control condition teaching about brain function but not its malleability. The central, preregistered outcome was grade point average (GPA) in core ninth-grade classes (0–4.3 scale), reflecting coursework up to 8–9 months after receiving the intervention.

Diverging Interpretations

Not surprisingly, proponents and critics of mindset interventions interpret the study's results in substantially different ways (https://www.scientificamerican.com/article/debate-arises-over-teaching-growth-mindsets-to-motivate-students/). The study authors emphasize real-world utility, but many critics dismiss the results based on the 'small' overall treatment effect estimate: 0.05 higher GPAs (P <0.001), also corresponding to 0.05 standard deviation (SD) units. This dismissal makes sense given two prevalent norms in psychological research: (i) a focus on estimating 'the effect', ignoring effect heterogeneity; and (ii) use of arbitrary, context-free effect size guidelines such as 0.20 SDs being a small effect [8]. Explaining some of the differing interpretations, the authors argued that both norms have limited utility in this context.

Regarding the first norm, estimating the aggregate effect was only tangential to the study's goals; the preregistration even explicitly hypothesized 'a very small (near zero)' average effect (https://osf.io/tn6g4). The authors instead focused on a key subgroup with theoretical and policy relevance: prior lower-achieving students. The treatment effect doubled to 0.10 GPA units for this subgroup. Hence, the intervention helped struggling students, even if it may not have created more academic superstars (at least in terms of GPA). Importantly, this subgroup analysis was not a 'fishing expedition' to find significant effects: it was preregistered prior to data analysis, building on prior evidence, including another large preregistered RCT [3]. Treatment effects for this focal subgroup also varied; they occasionally exceeded 0.20 GPA units (in 13% of schools based on heterogeneity estimates) and were stronger when peer norms aligned with the intervention's message.

Regarding the second norm (effect size guidelines), the authors argued that cost effectiveness and benchmark comparisons must be considered when evaluating the size of improvements, reflecting guidelines from applied fields such as education and public health. The intervention lasted less than 1 hour, required no teacher training, and is freely available online (https://www.perts.net/orientation/hg). Even more intensive educational interventions routinely fail to show large effects when rigorously evaluated at a national scale. A recent review of 59 RCTs commissioned by the US Department of Education via the National Center for Education Evaluation

(NCEE) found a median treatment effect of 0.05 SDs (Figure 1) [9]. These NCEE trials are relevant because, like the new mindset study, they evaluated the effectiveness of promising interventions in routine applied conditions, using rigorous experimental designs and preregistered analysis plans, usually focusing on distal, policy-relevant outcomes (e.g., grades, standardized test scores), and applying conservative analyses (e.g., including all randomized students, regardless of whether they received the full intervention).

As these sobering results show, promising effects from small-scale studies (e.g., controlled laboratory experiments) likely will not directly translate into improved success on key academic outcomes at a national scale. In this context, a reliable effect of 0.10 on ninth-grade core GPA for lower-achieving students is notable, with policy implications for reducing course failure rates.

Study Implications

Debate about this new national RCT therefore illustrates at least two broader points. First, investigating systematic effect heterogeneity is vital for understanding what works, for whom, and under what conditions. Second, interpreting quantitative effect sizes is not merely a statistical exercise; it also requires careful consideration of the study's context.

Obviously, the results are a far cry from eliminating achievement gaps and revolutionizing educational systems. Researchers and educators should not expect any single intervention to do so, especially for one lasting less than an hour. As mindset proponents have argued, these brief social psychological interventions 'will improve outcomes only when other aspects of the system necessary for improvement are in place' [6].

One serious concern, however, is whether applied audiences have heard or sufficiently understood these nuanced interpretations, given that nuances can often be stripped away, and misinterpretations can be likely when scientific research is filtered through mass media. Having unrealistically high expectations that 'equality happens' when creating growth mindset classrooms (https://www.ted.com/talks/carol_dweck_the_power_of_believing_that_you_can_improve) might steer educators and policymakers away from more comprehensive reform. At the same time, however, mindset scholars have also noted in practitioner journals that these interventions 'catalyze the effects of high-quality educational reforms, but don't replace them' [10].

Studying educators' beliefs about the promise of mindset interventions therefore could be one fruitful future research direction [4], along with many other interesting questions that the national study raises. What are the cognitive, social, and behavioral mechanisms explaining the improved grades months later in real-world settings? Why do some other large-scale RCTs [4] find different results? Does training students versus teachers make a difference? How should psychological and cognitive scientists conceptualize and empirically study effects less than 0.2 SDs in typical applied contexts? Investigating these questions could advance fundamental theory, scholarly debate, and applied insights for maximizing cognitive potential.

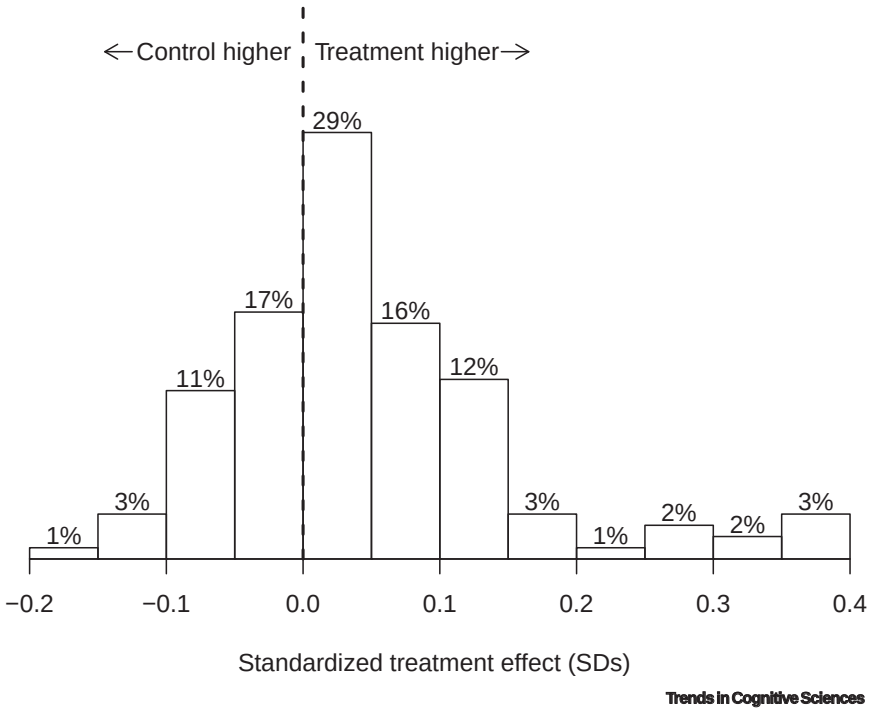

Figure 1. Distribution of Standardized Treatment Effects from 59 National Center for Education Evaluation Trials [9]. These large-scale trials, which evaluated promising educational interventions, share many methodological features in common with the national mindset study that might help explain the generally small statistical magnitudes (although additional considerations such as developmental stage may be important). These features include: (i) using rigorous experimental designs with active control groups, (ii) sampling from heterogeneous national populations, (iii) preregistering the analysis plan and outcome measures, (iv) examining effectiveness in routine conditions with minimal oversight from the intervention developers, and (v) commissioning independent evaluators to randomize participants and preprocess the data. Abbreviation: SD, standard deviation.

Acknowledgments

The author thanks Alice Eagly, Allison Master, David Yeager, and Sara Hart for comments on an

earlier version of the manuscript, as well as several commentators to a Twitter thread that initiated this article.

[1]American Institutes for Research, Research and Evaluation, 10 S. Riverside Plaza, Chicago, IL 60606, USA

*Correspondence:
dimiller@air.org (D.I. Miller).

https://doi.org/10.1016/j.tics.2019.08.005

References

1. Sisk, V.F. *et al.* (2018) To what extent and under which circumstances are growth mind-sets important to academic achievement? Two meta-analyses. *Psychol. Sci.* 29, 549–571
2. Paunesku, D. *et al.* (2015) Mind-set interventions are a scalable treatment for academic underachievement. *Psychol. Sci.* 26, 784–793
3. Yeager, D.S. *et al.* (2016) Using design thinking to improve psychological interventions: the case of the growth mindset during the transition to high school. *J. Ed. Psychol.* 108, 374–391
4. Foliano, F. *et al.* (2019) *Changing Mindsets: Effectiveness Trial*, National Institute of Economic and Social Research
5. Miller, D.I. and Halpern, D.F. (2014) The new science of cognitive sex differences. *Trends Cogn. Sci.* 18, 37–45
6. Walton, G.M. and Wilson, T.D. (2018) Wise interventions: psychological remedies for social and personal problems. *Psychol. Rev.* 125, 617–655
7. Yeager, D.S. *et al.* (2019) A national experiment reveals where a growth mindset improves achievement. *Nature* Published online August 7, 2019. https://doi.dx.org/10.1038/s41586-019-1466-y
8. Funder, D.C. and Ozer, D.J. (2019) Evaluating effect size in psychological research: sense and nonsense. *Adv. Methods Pract. Psychol. Sci.* 2, 156–168
9. Lortie-Forgues, H. and Inglis, M. (2019) Rigorous large-scale educational RCTs are often uninformative: should we be concerned? *Educ. Res.* 48, 158–166
10. Yeager, D.S. *et al.* (2013) Addressing achievement gaps with psychological interventions. *Kappan* 94, 62–65