psyche
know your self

What the marshmallow test got wrong about child psychology

by Tyler W Watts, developmental psychologist

Photo by Bill Anron/Getty Images

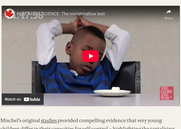
Tyler W Watts is an associate professor of psychology and education in the Department of Human Development at Teachers College, Columbia University. His research, which focuses on educational interventions, the worldwide as a research assistant professor at New York University for his postdoctoral training, and he received his PhD in Education from the University of California, Irvine. Across his projects, he studies how educational interventions designed to boost children's early cognitive and behavioural skills will make long-lasting changes on developmental trajectories.

Edited by **Christian Jarrett**

Published 4 August 2025

Self-control, grit, growth mindset – trendy skills won't transform children's lives, but more meaningful interventions can

A child squirms in her chair as she faces a single marshmallow. She's been told that if she can hold out for a little while – resist the temptation to take the treat – she'll be rewarded with two marshmallows as a prize for her self-restraint. The marshmallow test, first devised by Walter Mischel in the 1960s to measure a young child's ability to delay gratification, helped usher in a fascination with early childhood capacities and skills that has persisted to the present day – you may have seen some of the more rare videos on social media or YouTube showing children taking the test.



Mischel's original studies provided compelling evidence that very young children differ in their capacities for self-control – highlighting the tantalising possibility that key features of adult personality typically take root before a child has even turned five years-old. These studies eventually caught the imagination of the curious public when follow-up work in 1990 showed that children who waited longer during the marshmallow test were more successful as adolescents and adults, in terms of having higher SAT scores and being better-adjusted socially. This longitudinal work fuelled expectations that the early ability to delay gratification constituted a foundational life skill and, moreover, that interventions targeting this early capacity might unlock long-term benefits that would set children on higher-functioning trajectories into adulthood.

But in 2018, my colleagues and I muddied the picture. We published a re-examination of whether young children's performance on the marshmallow test really does predict how well they will fare in adulthood. The study gained some notoriety because we found that, although performance on the task was correlated with achievement in adolescence, this relationship all but vanished when we accounted for other important factors in a child's life, such as their general cognitive ability and socioeconomic status. In other words, we found that, if children were matched on these other factors, their ability to delay gratification was no longer related to their behavioural or academic outcomes later in life. We also recently followed up on this work and found the same was true for success in adulthood – once other factors are taken into account, the marshmallow test loses its predictive power.

Since we released our 2018 study, it just happened to be a peak moment of public interest in the 'replication crisis' in psychology when other famed Psych 101 studies were also receiving renewed critical attention. Unfortunately, Mischel's marshmallow test was swept up unfairly in a damning narrative: 'Another (old) psychology study disproved!' But it wasn't so much that we'd failed to replicate Mischel's famous longitudinal results. It was more that our study confirmed a nagging message that the field is resistant to accepting: *there are no individual skills that are silver bullets.*

The marshmallow test research held so much appeal because it tapped into what seems like an intuitively sensible approach. Adults are widely considered resistant to change and difficult to reach with interventions. You can't teach an old dog new tricks, as they say. In contrast, young children are considered highly malleable. If we could identify foundational skills in childhood that are susceptible to training, then perhaps we could design early interventions to target these key capacities and thereby produce outsize impacts on long-term functioning.

The marshmallow test research highlighted the ability to delay gratification as one of these foundational skills. Enterprising psychologists have since lined up to propose alternatives: 'grit', 'growth mindset', 'executive functioning', among others. The recipe for all these candidates is fairly straightforward. As with the marshmallow test, the research gains steam when correlational studies show that an early life skill is predictive of important later outcomes. Early executive functioning *predicts* later academic ability. *Grit predicts* educational attainment. Self-control in childhood *predicts* later economic success and health. Therefore, target these skills and you set children up for life, or so the logic goes.

> “The approach of finding unique skills with outsize impacts on adult functioning is largely misguided”

Psychologists play into this expectation when they float the possibility of interventions in their papers, books and public talks. Educational advocates, teachers and parents are desperate for solutions to hard problems regarding student underachievement and gaps in educational outcomes. If we can alter a child's entire developmental trajectory by simply helping them learn self-control and grit, why wouldn't we try?

Unfortunately, this cycle of hype-fuelled research has failed to find childhood interventions that can reliably generate desirable long-term effects. The reality is that intervention efforts often fail to make long-lasting impacts on targeted skills, even when those interventions target skills that were shown to be so predictive of later outcomes in correlational research. Instead, the research shows that when children are given an intervention for a specific psychological capacity, any advantages are usually short-lived, leaving the children in an apparently similar position to where they began. This problem, known as 'fadeout', has been the focus of our team's research over the past several years. Through this work, I have become increasingly convinced that the entire approach of finding unique skills that will carry outsize impacts on broad domains of adult functioning is largely misguided.

To illustrate why, let me describe our line of work on childhood mathematics ability – an early capacity that has inarguable importance for school success, though it lacks the *je ne sais quoi* of executive functioning or grit. Following the recipe described above, study after study had demonstrated the predictive validity of very early mathematical abilities for later success in school and life. In other words, children who are ahead of their peers in mathematics early in school tend to stay ahead of their peers across multiple subjects. Mathematics skills are also obviously malleable – kids learn them every day in school, and no one is born a mathematician. So, could an early intervention aimed at mathematics skills be a silver bullet to set children on the path to success?

In fact, although interventions targeting early mathematics ability led to substantial benefits, they were short-lived. Several years after the intervention ended, the children who received the intervention early in life were faring no better in mathematics than the children who did not get the intervention.

Why did we see fadeout, and why did the correlational research suggest that early mathematics ability would provide durable long-term benefits to academic trajectories? Although we are still trying to understand why fadeout occurs, our work has suggested that the story for mathematics is similar to why the marshmallow test lost its predictive power once broader child and environmental factors were taken into account.

> “They suggested that interventions targeting social-emotional skills might produce enduring effects”

The issue is that a child's mathematics ability, much like their ability to delay gratification, is more a proxy for broader life factors than the direct cause of them. Children who are more skilled at mathematics early in life also tend to be more skilled at mathematics later life because most of the factors that led them to be more advantaged early – home environments, resources, genes, parents, personality, cognitive ability (the list goes on) – are still influencing them later on. An intervention simply targeting mathematics, or any other psychological skill, likely has no effect on these other life forces that were present before the intervention started, and will continue to be present after the intervention ends.

This has been a hard lesson for the field to learn, and many psychologists and educators have argued that there must surely be some class of skills or educational interventions that will reliably produce effects resistant to fadeout. Indeed, many have suggested that interventions targeting 'non-cognitive skills', or social-emotional skills, such as self-control or motivation, might produce enduring effects. They've argued that perhaps the fadeout problem only applies to the kinds of cognitive skills directly taught in schools.

To shelter understand the breadth of the fadeout problem, our team started using meta-analytic approaches (ie, looking for patterns across many underlying studies) to examine skill fadeout and persistence across widely varying interventions. We compiled a dataset of approximately 85 randomised controlled trials of various educational programmes, all targeting a broad range of child and adolescent skills and competencies. Importantly, these studies tracked children into the long-term after the intervention ended, allowing us to observe how initial intervention-spurred benefits unfolded over time. We began by replicating our earlier work on mathematics achievement fadeout. Much like we had seen with our studies of early mathematics programmes, cognitive skill impacts on other domains, such as reading achievement, also faded by 40-50 per cent in the first year following the intervention. Surprisingly, we found a similar fadeout pattern for the interventions targeting social-emotional skills and capacities.

So where does this leave us? Should we give up all hope that educational interventions can have enduring effects? No. The literature clearly supports the idea that high-quality long-term educational programmes can affect key outcomes in adulthood, especially for children who come from disadvantaged communities. The issue is that we still lack an understanding of the skills and capacities that reliably lead to these longer-term effects.

> “The insight that silver bullets are unlikely to exist is not new, nor should it be controversial”

Some epistemic humility seems appropriate here, and psychologists who continue to use correlational research to make claims about 'silver bullet' skills are not helping us move forward. If fadeout tends to occur for cognitive and social-emotional skills alike, then it seems unlikely that we will ever find a single skill or capacity that is key to adult outcomes. Instead, we should probably face up to the fact that developmental trajectories are shaped by myriad factors that are highly complex and difficult to predict, and any single skill is likely to be only a very small piece of a much broader set of factors that are shaping human development over time.

Indeed, the intervention studies that have provided the best evidence for long-term effects, such as the Carolina Abecedarian Project conducted in North Carolina from 1972-82, did not target silver bullet skills. Instead, they attempted to generate fundamental change for children by improving their environmental experiences in broad and ambitious ways that were resource-intensive but meaningful (the Abecedarian Project provided intensive, full-day centre-based care from birth through age five, combined with comprehensive medical services, nutrition support and family counselling). So, the next time you hear about a flashy 'new' skill that has the potential to change everything, you should react with a healthy dose of scepticism that psychology is unlikely to produce easy answers. Perhaps this insight should also bring a sense of relief, as the promise of 'hacking' our psychology for short-cut benefits also implies that human potential is ever-reflant on specific skills that many people may be unlikely to ever acquire.

Of course, the insight that silver bullets are unlikely to exist is not new, nor should it be controversial. More than 100 years ago, Edward Thorndike debunked claims that learning Latin would have outsize effects on other cognitive capacities, despite the appealing theory that Latin provides foundational thinking skills that should help in many domains of reasoning. Yet, here we are more than a century later still in search of the holy grail. Let's hope that we start to pay attention to our own evidence and not waste another 100 years continuing the search.

FOCUS AND ATTENTION CHILDHOOD AND ADOLESCENCE LEARNING AND EDUCATION