

**Abstract**

Children developing in harsh conditions often perform poorly on standard cognitive tests. However, the hidden talents approach proposes that performance for some abilities may be enhanced by adversity or improved under ecologically relevant conditions. We examine ecologically relevant content in cognitive testing. Sampling 618 socioeconomically diverse adolescents, we measured performance on two versions of an attention-shifting and working memory updating task—one abstract, the other ecological. Across multiple sources, we measured environmental unpredictability, violence, and poverty, and tested adversity  $\times$  task version interactions using a preregistered multiverse approach. We did not find adversity  $\times$  task version interactions for attention shifting, but there were main effects of unpredictability and violence, indicating improved attention shifting irrespective of task version. For working memory updating, violence- and poverty-exposed youth performed worse on the abstract but not the ecological version, equalizing performance among adversity-exposed youth. We found no relation between updating performance and unpredictability.

*Keywords:* hidden talents, working memory updating, attention shifting, adversity, adaptations to stress

**Hidden talents in context: Can ecologically relevant stimuli improve cognitive performance among adversity-exposed youth?**

Children living in harsh conditions tend to perform poorly on many cognitive tests. For example, children raised in poverty often show difficulties learning in school and score lower on standardized tests<sup>1-4</sup>. Harsh and unpredictable environments are thought to cause high levels of chronic stress, which impacts brain regions responsible for memory and learning<sup>5</sup>. Thus, the general consensus is that exposure to adversity undermines cognitive development and leads to a host of cognitive impairments<sup>4</sup>. Although it is clear that harsh environments are harmful to children, there is growing interest in the *hidden talents approach*, which proposes that some abilities may be enhanced by adversity, especially those abilities that are ecologically relevant to lived experience. Thus, hidden talents are skills and abilities that help children meet real-world challenges in adverse environments<sup>6-9</sup>.

The hidden talents framework assumes that development shapes the skills and abilities of individuals in relation to their environment<sup>6-9</sup>. Accordingly, exposure to adversity should shape skills and abilities well-suited for dealing with the challenges associated with harsh environments. However, because specific dimensions of adversity (violence versus resource-scarcity) may pose unique challenges<sup>7,9-11</sup>, the hidden talents approach proposes that cognitive skills should match specific dimensions of adversity<sup>9</sup>. For instance, physically abused children can detect angry facial expressions quickly and accurately<sup>12-14</sup>, people living in poverty have shown some evidence of enhanced social perception<sup>15</sup>, and anxiously attached people are good at deception detection<sup>16,17</sup> (see<sup>6,7</sup> for reviews).

Although some abilities themselves are more ecologically relevant to specific dimensions of adversity, hidden talents may also emerge under ecologically relevant conditions. For example, previous studies of hidden talents have tested cognition under ecologically relevant *contexts and situations*. In particular, studies have experimentally manipulated the current context to be uncertain and mildly stressful (or not) and measured performance on executive functioning and working memory tasks<sup>18,19</sup> (but see<sup>20</sup> for similar results without uncertainty). Strikingly, people from unpredictable environments showed improved performance on attention-shifting and working memory updating tasks. However, this effect only emerged under conditions of uncertainty. There was no performance improvement under control conditions.

Hidden talents may also emerge when testing *content and stimuli* are ecologically relevant to adversity-exposed individuals. Typical cognitive tests use abstract content that is largely detached from the real world. Such abstract materials may disadvantage youth with limited exposure to formal education. Cultural psychologists and anthropologists have demonstrated this fact across the world. For example, people from low

income and non-western cultures tend to perform worse on standard versions of cognitive tests, but this performance gap closes or disappears with more ecologically relevant versions of the same tests<sup>21–25</sup>. These findings pose an important question: Can ecologically relevant testing materials improve cognitive performance among adversity-exposed youth?

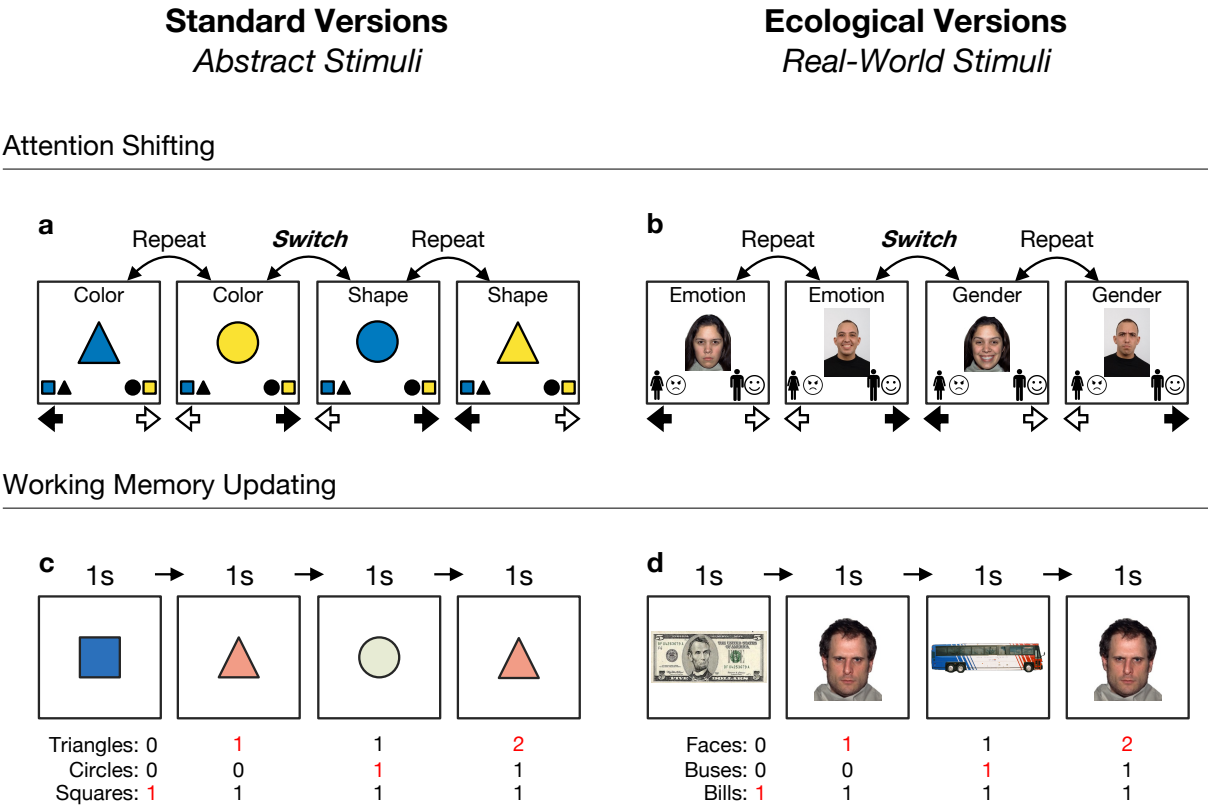
One way to test this question is to simply replace sterile testing materials with ecologically relevant content. For example, one study tested memory and reasoning among young adults who varied in their exposures to violence<sup>26</sup>. They created two versions of a memory and reasoning task, one neutral (chronological age) and one more ecologically relevant (social dominance). They found that people with more current exposures to violence were equally good, or even better, at memorizing social-dominance relationships compared to people with fewer exposures. Notably, such violence exposures were associated with worse memory for neutral (age) relationships. There were no associations revealed between violence exposures and reasoning performance. Overall, this study suggests that ecologically relevant testing materials may “level the playing field” for adversity-exposed youth<sup>7,9,27,28</sup>. In addition, although executive functions and working memory are related to many educational outcomes, it is unclear how schools and teachers might leverage skills that specifically depend on inducing uncertainty. That seems impractical, at best, and unethical, at worst.

Here we address this issue by evaluating the role of ecologically relevant content in attention-shifting and working memory tasks. To do so, we sampled 618 adolescents from a socioeconomically diverse population. Using interviews, self-reports, and school records, we measured exposure to environmental unpredictability, violence, and poverty. We then tested performance on two widely used tasks: attention shifting and working memory updating. There were two versions of each task. The first used standard abstract stimuli and the second replaced these stimuli with more ecologically relevant ones (see Figure 1). We then tested the interactive effect of task content and adversity exposure.

We expected adversity-exposed adolescents to benefit more from ecologically relevant stimuli (compared to abstract stimuli) than their counterparts. As such, our primary analyses involved two preregistered confirmatory analyses and one exploratory analysis. Our confirmatory analyses focused on environmental unpredictability and violence exposure because both dimensions have been associated with improved performance under ecologically relevant conditions (contexts and testing materials) and represent specific environmental challenges. Our exploratory analyses focused on exposure to poverty, which is both relevant to hidden talents (e.g., hidden talents could develop in response to specific challenges of poverty) and a general proxy to adversity rather than a specific exposure.

We also evaluated the relative robustness of our models to arbitrary data processing decisions using a multiverse analysis approach<sup>29,30</sup>. Specifically, we systematically and explicitly generated all possible

combinations of (arbitrary) data processing decisions and analyzed them simultaneously. We conducted a separate multiverse analysis for each task (attention shifting and working memory updating) and each measure of adversity exposure combination. All multiverse analyses used the same underlying statistical models.



the interaction between each adversity dimension with task version (abstract vs. ecological) on attention-shifting and working memory performance. We ran separate analyses for each adversity dimension and outcome combination.

Before conducting analyses, we identified six arbitrary data processing decisions each with two alternatives (see [Arbitrary Decisions](#) section). These criteria were 1) including the after-school club sample or not; 2) including participants with minor distractions during testing or not; 3) including participants who missed one trap question or not (those who missed more than 1 were already excluded); 4) including participants receiving special education (participants receiving more > 60 minutes/day were already excluded) or not; 5) including participants who performed at or below chance levels of accuracy on the attention-shifting task or not (see [attention shifting data cleaning procedure](#) for more information); and 6) residualizing social desirability scores out of adversity scores or not. Using all possible combinations of arbitrary decisions, we constructed a multiverse of 64 datasets. The minimum and maximum sample sizes were 418 and 618, respectively.

We ran the same underlying mixed-effects linear regression model. Analyses were conducted using R<sup>31</sup> and mixed models were fit using the lme4 package<sup>32</sup>. Our mixed-effects models included five fixed terms: the main effect of task version (ecological or abstract version, within-subjects), the main effect of childhood adversity (between-subjects), the interaction between task version and adversity exposure, the main effect of age (control variable), and the main effect of test distraction ratings (control variable). Task version was grand mean centered (abstract versions = -1, ecological versions = 1). All childhood adversity measures were standardized before model fitting, and all models included a random intercept for task version nested in participants.

Our models for attention shifting and working memory updating differed in one respect. Our sample included a small subset of siblings ( $n = 93$ , 45 families). All sibling groups were pairs of two (43 sibling pairs) except for 1 group of three and 1 group of four siblings. We computed intraclass correlations (ICCs) for sibling groups for attention-shifting and working memory updating scores. ICCs for attention shifting were essentially zero, but they were non-zero for working memory updating scores ( $ICC_{standard} = 0.02$ ,  $ICC_{ecological} = 0.24$ ). Thus, all models analyzing working memory updating included an additional random intercept for siblings. Non-siblings were given the same sibling-group code, so the model could compute an intercept for non-siblings.

To facilitate interpretation of multiverse results, we report median standardized  $\beta$ s, median 95% confidence intervals, and percent significant for each term in our model (see Table 2). Percent significant refers to the number of analyses (out of 64) that resulted in a significant effect for a particular term and reflects the

relative sensitivity (or robustness) of an effect to arbitrary data processing decisions. For the main effects of adversity and task version  $\times$  adversity interactions, we used a bootstrapping technique<sup>30,33,34</sup> to compute overall  $p$ -values, which indicate whether a  $\beta$  of a term is larger than expected given the null (the median  $\beta = 0$ ; see [Statistical Inferences](#)).

Table 1 displays bivariate associations between all composite measures of adversity, including components of each, and descriptive statistics. Our primary interests were the main effect of adversity and the interaction of adversity with task-version. Table 2 shows median effect sizes ( $\beta$ -coefficients), sample sizes, and bootstrapped  $p$ -values for attention-shifting and working memory updating analyses for the main effect of adversity (unpredictability, violence, and poverty) and the interaction term (task version  $\times$  adversity). Table 3 reports median simple effects of task version across high and low levels of adversity and median simple effects of adversity across abstract and ecologically relevant task versions.

Figures 2 and 3 visualize multiverse task version  $\times$  adversity interaction results for attention shifting and working memory performance, including depictions of performance on abstract and ecological task versions across high ( $+1 SD$ ) and low ( $-1 SD$ ) adversity exposure (Figures 2a and 3a),  $p$ -curves associated with each interaction term (Figures 2b and 3b), sorted interaction  $\beta$ -coefficients across arbitrary decisions (Figures 2c and 3c), sample sizes (Figures 2d and 3d), and a specification grid showing the arbitrary data processing decisions associated with each effect (Figures 2e and 3e).

We preregistered all primary analyses on the Open Science Framework (OSF; see the [osf project link](#), including an [updated primary analysis plan](#), which included more details. We conducted secondary analyses after our primary analyses (see [Secondary Analyses](#) below). Before conducting these analyses, we uploaded secondary analysis plans (see the [first secondary analysis plan](#) and the [second plan](#), explaining the goals and purpose of each secondary analysis. A Red Team Critic<sup>35</sup> successfully reproduced and verified (see this [red team report](#)) all analyses prior to submission.

## Primary Analyses

### *Attention Shifting*

Across each set of multiverse analyses, there were significant main effects of task version. All median effects sizes for task version were  $\beta_s = 0.33$  (all  $ps < .05$ ) for multiverse models testing unpredictability, violence exposure, and poverty exposure, indicating that switch costs (see [Attention Shifting](#) for switch cost explanation) were moderately higher for the ecological version compared to the abstract version (see SI Figures 1-3). That is, reaction times on switch trials, compared to repeat trials, were longer when stimuli

**Table 1**

*Bivariate correlations and descriptive statistics for all measures of adversity and their subcomponents.*

Variable	Unpredictability			Violence			Poverty			
	1	2	3	4	5	6	7	8	9	10
1. Perceived	-	601	601	601	601	601	601	599	558	601
2. Interview	0.50**	-	618	601	601	601	601	616	574	618
3. Composite	0.87**	0.87**	-	601	601	601	601	616	574	618
4. Neigh.	0.40**	0.20**	0.35**	-	601	601	601	599	558	601
5. Fighting	0.25**	0.10*	0.20**	0.60**	-	601	601	599	558	601
6. Composite	0.36**	0.17**	0.31**	0.89**	0.89**	-	601	599	558	601
7. Perceived	0.45**	0.28**	0.42**	0.44**	0.34**	0.44**	-	599	558	601
8. Interview	0.13**	0.15**	0.16**	0.18**	0.12**	0.17**	0.23**	-	573	616
9. School	0.18**	0.24**	0.24**	0.15**	0.12**	0.15**	0.28**	0.36**	-	574
10. Composite	0.37**	0.29**	0.37**	0.38**	0.28**	0.37**	0.75**	0.77**	0.62**	-
Min	1.00	-0.79	-1.16	1.00	1.00	-1.09	1.00	-2.06	-0.50	-1.47
Median	1.83	-0.39	-0.16	1.71	2.00	-0.25	2.00	0.05	-0.50	-0.06
Max	4.67	2.69	3.34	4.57	8.00	3.64	4.67	1.96	0.50	2.08
Range	3.67	3.48	4.50	3.57	7.00	4.73	3.67	4.03	1.00	3.54
Mean	1.99	-0.00	0.01	1.94	2.31	-0.00	2.20	0.01	-0.05	-0.01
SD	0.73	0.83	0.87	0.75	1.43	0.89	0.67	0.87	0.50	0.62

*Note:* Column headers denote adversity constructs and rows indicate components. The upper triangle contain sample sizes for each correlation. \*  $p < .05$ , \*\*  $p < .01$

**Table 2**

*Median standardized effects, 95% confidence intervals, percent significant, and bootstrapped p-values for the main effect of adversity and interaction terms for adversity by task version for attention shifting and working memory updating.*

	Main Effect				Interaction			
	$\beta$	95% CI	$p$ (%)	$p$	$\beta$	95% CI	$p$ (%)	$p$
<b>Attention Shifting</b>								
Unpredictability	-0.05	[-0.12, 0.01]	25.00%	0.022	0.02	[-0.03, 0.08]	0.00%	0.218
Violence	-0.06	[-0.12, 0.00]	21.88%	0.020	-0.02	[-0.08, 0.04]	0.00%	0.176
Poverty	-0.03	[0.03, -0.10]	0.00%	0.310	0.02	[0.07, -0.04]	0.00%	0.312
<b>Working Memory Updating</b>								
Unpredictability	-0.04	[-0.11, 0.04]	0.00%	0.106	0.02	[-0.03, 0.06]	0.00%	0.374
Violence	-0.06	[-0.13, 0.02]	18.75%	0.082	0.04	[-0.00, 0.09]	40.62%	0.036
Poverty	-0.10	[-0.03, -0.18]	100.00%	0.000	0.05	[0.09, 0.00]	53.12%	0.012

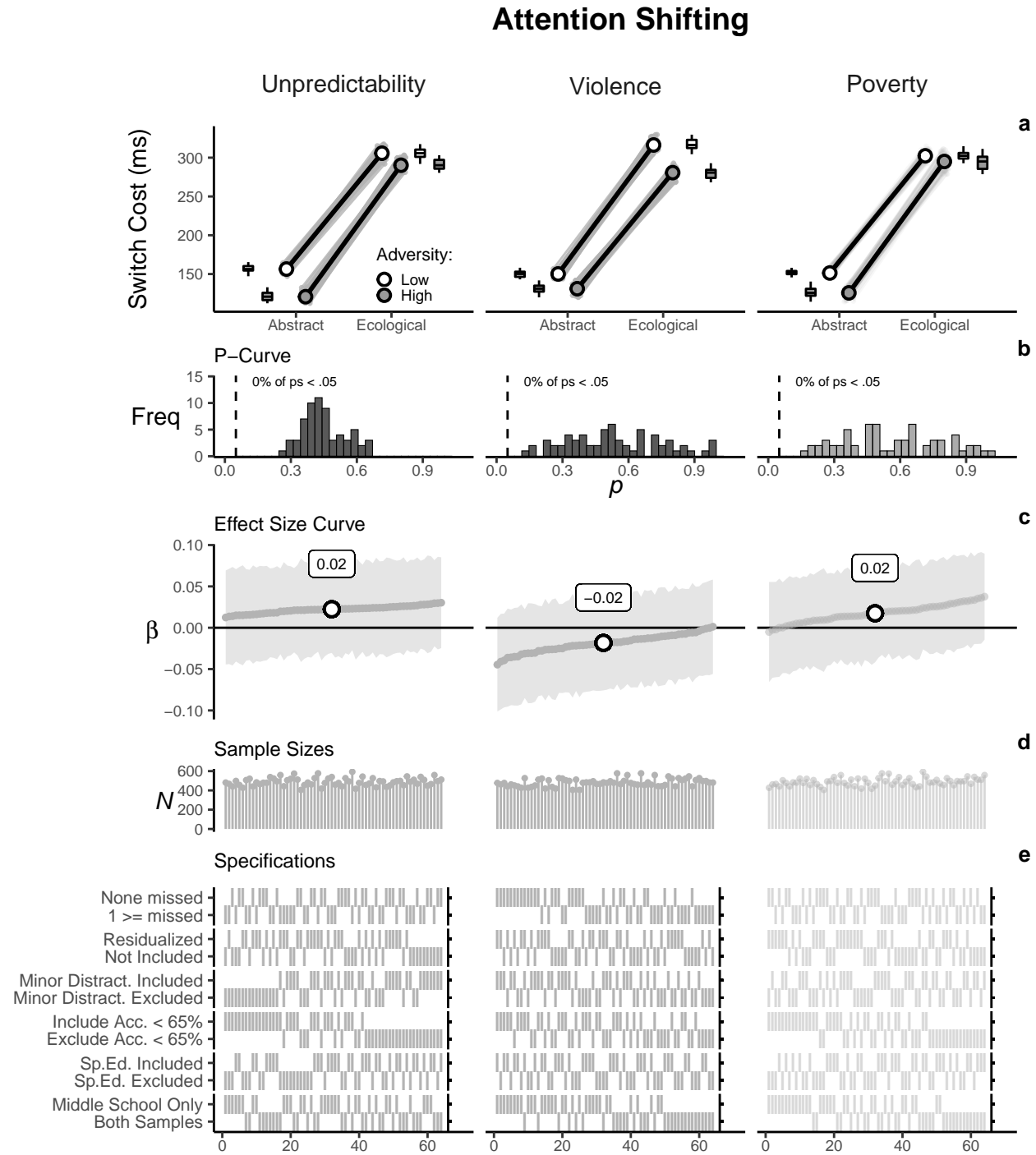
*Note:*

The  $p$  (%) column reflects the number of analyses that produced  $p$ -values  $< .05$  out of all analyses for a given multiverse analyses. The total number of analyses for each measure was 64. Overall  $p$ -values were computed using a bootstrapping resampling method and reflect the probability of obtaining an effect size as extreme or more extreme given the median effect is 0.

were ecologically relevant as opposed to abstract. However, there were no effects of age (all median  $\beta$ s = 0.02, all  $ps > .05$ ) or test distraction ratings (all median  $\beta = 0.02$ , all  $ps > .05$ ) on overall switch costs (see SI Figures 1-3).

Overall, there were no significant interaction effects between task version and unpredictability, violence exposure, or poverty exposure (see Table 2, Table 3, and Figure 2b). However, analyses produced main effects of unpredictability (median  $\beta = -0.05$ , 25% of  $ps < .05$ , overall  $p = 0.022$ ) and violence exposure (median  $\beta = -0.06$ , 21.88% of  $ps < .05$ , overall  $p = 0.020$ ). (See Table 2.) There were no main effects of poverty exposure (median  $\beta = -0.03$ , no  $ps < .05$ , overall  $p = 0.310$ ); see Table 2). That is, exposure to more unpredictability and violence was associated with smaller switch costs, suggesting such exposures were associated with somewhat better shifting performance on average, irrespective of whether stimuli were abstract or ecologically relevant.



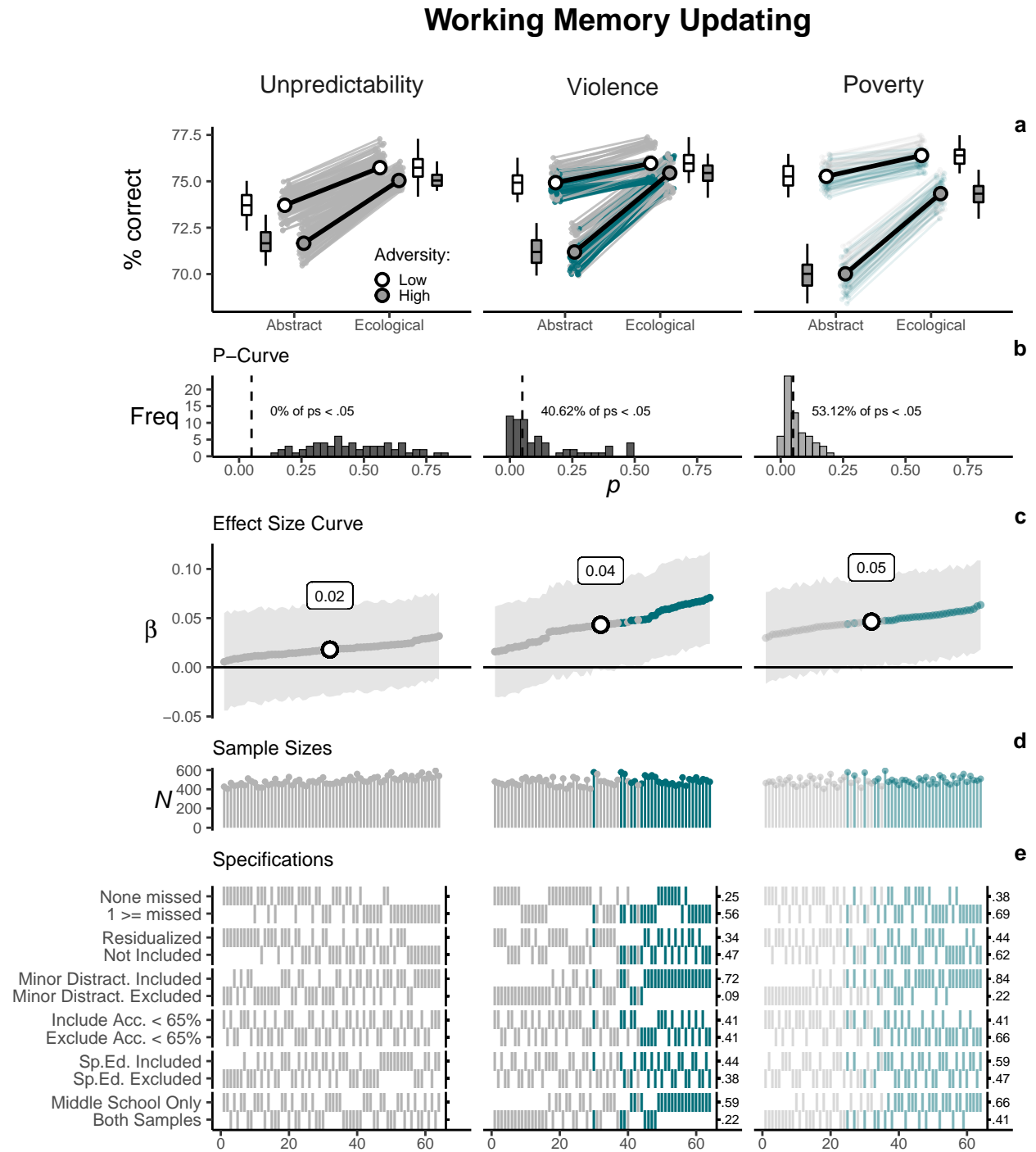


**Figure 2.** Visualization of multiverse attention shifting results. Specifically, a) visualizes the multiverse task version  $\times$  adversity interaction on abstract and ecological task versions across high ( $+1$  SD) and low ( $-1$  SD) adversity exposure, b) plots  $p$ -curves associated with each interaction term, c) plots sorted interaction  $\beta$ -coefficients across each arbitrary decision, d) plots the sample sizes for each effect, and e) is a specification grid indicating the data processing decisions associated with each effect.

### Working Memory Updating

Across each set of multiverse analyses, there were again significant main effects of task version. All median effects sizes for task version were  $\beta = 0.08$  (all  $ps < .05$ ) for models testing unpredictability, violence exposure, and poverty exposure, indicating that updating performance was better when stimuli were ecological relevant compared to abstract, on average (see SI Figures 4-6). There were also consistent main effects of age; median effects were  $\beta = 0.14$  (all  $ps < .05$ ) for unpredictability models,  $\beta = 0.15$  (all  $ps < .05$ ) for violence exposure models, and  $\beta = 0.15$  (all  $ps < .05$ ) for poverty exposure models (see SI Figures 4-6). That is, older adolescents scored higher on working memory updating, on average. Finally, there were some main effects of test environment distraction ratings. Median effects sizes were  $\beta = -0.1$  (62.5% of  $ps < .05$ ) for unpredictability models,  $\beta = -0.11$  (75% of  $ps < .05$ ) for violence exposure models, and  $\beta = -0.1$  (62.5% of  $ps < .05$ ) for poverty exposure models, indicating that more test distractions reduced working memory updating performance, on average (see SI Figures 4-6).

Multiverse analyses revealed a main effect of poverty exposure (median  $\beta = -0.10$ , all  $ps < .05$ , bootstrapped  $p < .001$ ) but no main effects of unpredictability (median  $\beta = -0.04$ , no  $ps < .05$ , bootstrapped  $p = 0.106$ ) or violence exposure (median  $\beta = -0.06$ , 18.75% of  $ps < .05$ , bootstrapped  $p = 0.082$ ). Analyses also revealed two significant interactions: task version  $\times$  violence exposure (median  $\beta = 0.04$ , 40.62% of  $ps < .05$ , bootstrapped  $p = 0.036$ ) and task version  $\times$  poverty exposure (median  $\beta = 0.05$ , 53.12% of  $ps < .05$ , bootstrapped  $p = 0.012$ ). Adolescents exposed to low levels of violence or poverty tended to perform better on the abstract working memory updating task (violence exposure median  $b_{abstract} = -1.80$ , 81.25% of  $ps < .05$ ; poverty exposure median  $b_{abstract} = -2.66$ , all  $ps < .05$ ), compared to adolescents exposed to high levels of violence and poverty (see Table 3 and Figure 3). However, this performance gap closed on the ecological working memory task (violence exposure median  $b_{ecological} = -0.31$ , no  $ps < .05$ ; poverty exposure median  $b_{ecological} = -1.03$ , 7.81% of  $ps < .05$ ). Put differently, adolescents exposed to high levels of violence or poverty (+ 1 *SD*) performed better on the ecologically relevant version compared to their own performance on the abstract version (median simple  $b_{high\ violence} = 2.08$ , all  $ps < .05$ ; median simple  $b_{high\ poverty} = 2.12$ , all  $ps < .05$ ), whereas adolescents exposed to low violence or poverty (− 1 *SD*) performed similarly across task versions (median simple  $b_{low\ violence} = 0.58$ , no  $ps < .05$ ; median simple  $b_{low\ poverty} = 0.56$ , no  $ps < .05$ ). In this sense, ecologically relevant stimuli appeared to equalize working memory updating performance among adolescents exposed to high levels of violence and poverty (see Table 3 and Figure 3a).



**Figure 3.** Visualization of multiverse working memory updating results. Specifically, a) visualizes the multiverse task version  $\times$  adversity interaction on abstract and ecological task versions across high ( $+1$   $SD$ ) and low ( $-1$   $SD$ ) adversity exposure, b) plots  $p$ -curves associated with each interaction term, c) plots sorted interaction  $\beta$ -coefficients across each arbitrary decision, d) plots the sample sizes for each effect, and e) is a specification grid indicating the data processing decisions associated with each effect. Proportions of each arbitrary decision with  $p$ -values  $< .05$  are indicated on the right side of each specification grid. Blank proportions indicate proportions = 0.

## Secondary Analyses

We followed-up our primary analyses with two sets of secondary, exploratory analyses. We uploaded secondary analysis plans to the OSF prior to conducting them (see the [first secondary analysis plan](#) and the [second plan](#)). These analyses focused on two questions. First, in contrast to prior work<sup>18,19</sup>, we found limited support for improved working memory updating performance in relation to unpredictability (although we found a main effect of unpredictability on attention shifting). However, prior work measured retrospective perceptions of unpredictability whereas the current combined perceptions with more objective, interview-based measures. Thus, one possibility is that specific components of our composite adversity measures (unpredictability, violence exposure, and poverty exposure; see [Measures](#) section) have unique associations with attention shifting and working memory updating performance. To address this, we expanded our multiverse approach to include each component of our predictor variables as additional arbitrary decisions.

The second question focused on the interactive effects of task version  $\times$  violence exposure and task version  $\times$  poverty exposure on working memory updating performance. Specifically, primary analyses revealed an equalization effect for both violence and poverty exposures on working memory updating. However, because measures of violence and poverty were not included in the same statistical model, it is unclear whether these interactions are independent. For example, one possibility is that exposures to violence and poverty are independently associated with working memory updating performance. Alternatively, they may overlap, reflecting the same association. To test this question, we ran a focused multiverse analysis with the same controls, main effects of task version, poverty and violence exposure, and two interaction terms: task-version  $\times$  violence exposure and task-version  $\times$  poverty in the same model. If violence and poverty exposures interact independently with task version, both interaction terms should remain significant. However, if they represent overlapping effects on working memory updating performance, one or the other may drop out of significance, which would suggest that one or the other adversity dimension is driving the effects of the other.

## *Exploring Components of Adversity Dimensions*

We ran an expanded multiverse analysis to explore the effects of each component of our adversity composites (see [Measures](#) section) on attention shifting and working memory updating performance. Environmental unpredictability contained two components, perceived unpredictability (self-reported) and an interview-based measure tabulating family disruptions, parental figures, and residential moves. Violence exposure contained two self-report components: neighborhood violence and exposure to/involvement in vi-

**Table 3**

*Median simple effects (unstandardized) and percent significant for task version and adversity for attention shifting and working memory updating.*

	Task Version				Adversity Exposure			
	<i>b</i> (abs)	<i>p</i> (%)	<i>b</i> (eco)	<i>p</i> (%)	<i>b</i> (low)	<i>p</i> (%)	<i>b</i> (high)	<i>p</i> (%)
<b>Attention Shifting</b>								
Unpredictability	-18.22	10.94%	-7.29	0.00%	74.14	100.00%	84.88	100.00%
Violence	-9.62	0.00%	-18.61	25.00%	83.60	100.00%	74.34	100.00%
Poverty	-12.51	0.00%	-4.08	0.00%	75.17	100.00%	83.63	100.00%
<b>Working Memory Updating</b>								
Unpredictability	-0.95	0.00%	-0.32	0.00%	1.01	28.12%	1.68	100.00%
Violence	-1.80	81.25%	-0.31	0.00%	0.58	0.00%	2.08	100.00%
Poverty	-2.66	100.00%	-1.03	7.81%	0.56	0.00%	2.12	100.00%

*Note:* Simple effects for Task Version reflect the median effect of adversity in for the abstract version (abs) and ecological version (eco). Simple effects for Adversity reflect the effect of task version when adversity-exposure is low (-1 SD) and high (+1 SD). P-values < .05 out of all analyses for a given multi-verse analyses. The total number of analyses for each measure was 64.

olence. Poverty exposure contained three components: parental education/occupation (interview), perceived resource scarcity (self-report), and school-coded economic disadvantage (fee waivers, free- or reduced-priced lunch, and homelessness).

For this expanded multiverse approach, we retained all original data processing decisions (i.e., original 64 combinations of decisions) and added adversity component as an additional decision. For unpredictability, we analyzed each unpredictability component and the composite measure (composite analyses are redundant with primary analyses but were retained for comparison purposes), which resulted in 384 analyses (64 original  $\times$  three versions of unpredictability  $\times$  two performance measures [attention shifting and working memory]). We applied the same approach to violence exposure (three violence variables, 384 analyses) and poverty exposure (four poverty variables, 512 analyses). The full results of these analyses are reported in the supplement (see SI Table 2).

For attention shifting, consistent with our primary analyses, there were no interactions for any component of adversity and task version (see SI Table 2 and SI Figures 7-9). However, there were a few main effects of adversity. First, primary analyses revealed, under some analytic decisions, a main effect of unpredictability. Upon inspection of each unpredictability component, this effect appeared to be entirely driven by self-reported perceived unpredictability. Specifically, no main effect  $p$ -values for interview-based unpredictability were  $< .05$ , whereas 90.62% were for perceived unpredictability (see SI Table 2, SI Table 3, and SI Figure 7). In addition, violence exposure revealed that our primary analyses were driven by neighborhood violence (65.62% of  $ps < .05$ ) and not by direct witnessing or involvement in fights (no  $ps < .05$ , see SI Table 2, SI Table 3, and SI Figure 8). There were no main effects of poverty exposure components on attention shifting (see SI Table 2, SI Table 3, and SI Figure 9).

For working memory updating, there were no interaction between any unpredictability component (perceived or interview-based) and task-version (see SI Figure 10). However, our expanded analysis with violence exposure revealed that interactions obtained with the composite measure were not driven by neighborhood violence but instead by witnessing or involvement in violence (75% of interaction  $ps < .05$ , see SI Table 2, SI Table 3, and SI Figure 11). No interactions were significant with the neighborhood violence component of the violence composite. For poverty exposure, interview-based parental occupational prestige and school-reported economic disadvantage showed some interaction effects (parental education/occupational prestige 26.56% of  $ps < .05$ ; school coded economic disadvantage 18.75% of  $ps < .05$ , see SI Table 2, SI Table 3, and SI Figure 12). Self-reported perceived resource scarcity showed no interaction effects (see SI Table 2, SI Table 3, and SI Figure 12).

### *Comparing Violence and Poverty Exposure Interactions with Task-Version*

Our next secondary analysis was designed to compare the violence exposure  $\times$  task-version interaction with poverty exposure  $\times$  task-version interaction on working memory updating performance in the same model. To do so, we ran a multiverse analysis using the original set of 64 data processing decisions. We entered the same controls, main effects of task-version, poverty and violence exposure, and two interaction terms: task-version  $\times$  violence exposure and task-version  $\times$  poverty exposure. Analyses revealed a main effect of poverty (90.62% of  $ps < .05$ ); consistent with the primary analyses, lower poverty was associated with better updating performance under most analytic decisions. However, there were no main effects of violence exposure (no  $ps < .05$ ). For interactions, there were no effects of task version  $\times$  poverty (no  $ps < .05$ , see SI Figure 13) and only a small number of significant task version  $\times$  violence exposure effects (14.06% of  $ps < .05$ , see SI Figure 13). More specifically, when both interactions were entered into the same model, only task-version  $\times$  violence exposure remained significant, although only under a small set of analytic decisions.

Violence and poverty exposure interaction patterns were qualitatively similar to those obtained in our primary analyses (see Figure 3a and SI Figure 13). However, these secondary results suggest that the violence exposure interaction with task-version may be driving our poverty  $\times$  task version results for working memory updating. One possibility is that the shared variance between poverty and violence exposures ( $r = -0.37$ ) is responsible for both interaction effects. In other words, although poverty and violence exposure are theoretically distinct, our poverty composite may capture poverty-related violence exposures (e.g., high poverty neighborhoods contain more crime) and therefore interact with task-version. When directly compared, explicit measures of violence exposure remained significant (although only under a small set of analytical decisions), whereas poverty interactions dropped out entirely.

## **Discussion**

We examined how ecologically relevant testing materials affect cognitive performance among adversity-exposed youth. We measured three dimensions of adversity — environmental unpredictability, violence exposure, and poverty — and tested performance on attention-shifting and working memory updating tasks. We administered two versions of each task: one using abstract stimuli and the other using ecologically relevant stimuli, allowing us to compare performance within- and between-subjects.

Leveraging multiverse analysis<sup>29,30</sup>, we tested interactions between dimensions of adversity and task version on attention shifting and working memory updating performance. We did not find interaction ef-

fects for attention shifting. However, youth exposed to unpredictability and violence showed smaller switch costs on average, suggesting slightly improved attention shifting performance irrespective of task content. These findings partially diverge from prior work<sup>18,20</sup>. For example, whereas we found unpredictability was associated with improved attention shifting on average, Mittal and colleagues<sup>18</sup> found no difference in performance under neutral conditions. And, although they improved attention shifting for those exposed to high unpredictability under uncertainty, it remains unclear how this effect relates to those obtained in the current study. In addition, Fields and colleagues<sup>36</sup> found caregiving instability was associated with improved attention shifting, but Nweze and colleagues<sup>20</sup> found no association between adversity and attention shifting. However, they tested institutionalized children rather than measuring adversity dimensions<sup>20</sup>. In general, we should interpret cautiously. Although we followed field standards for calculating switch costs by computing difference scores, such scores ignore participant-level reaction times. Mixed-modeling approaches can account for both overall reaction times and differences across switch and repeat trials, but require more statistical power to test three-way interactions. Future research should consider more appropriate modeling approaches.

For working memory updating, we found lowered performance among violence- and poverty-exposed youth on the abstract working memory updating task compared to peers with reduced exposures. However, ecological materials produced an equalization effect: the updating performance gap between low and high adversity-exposed youth narrowed on the ecological updating task. Interestingly, updating performance was unrelated to unpredictability. These findings are partially consistent with prior work on adversity exposure and working memory updating performance<sup>19,20</sup>. For example, both Young and colleagues<sup>19</sup> and the current study found lowered abstract working memory performance among high adversity-exposed individuals. Although performance with ecological stimuli is not directly comparable to performance under manipulated uncertainty, both studies reduced the updating performance gap under ecologically relevant conditions. However, there are also discrepancies. For example, Young and colleagues<sup>19</sup> found effects with unpredictability and not poverty, whereas the current study found associations with violence and poverty but not unpredictability. Finally, Nweze and colleagues<sup>20</sup> found improved working memory performance among institutionalized youth without ecological stimuli or uncertainty manipulations. However, both their adversity and working memory measures were different. Such inconsistencies could be attributed to differences in sampling (adults vs. adolescents), the sampled range of adversity exposure (restricted vs. broad), or methods/design (within- vs. between-subjects design; lab vs. community/school settings). Future research should address these inconsistencies.

Our first set of secondary analyses focused on the components of our adversity composites. For at-



tention shifting performance, the main effect of our unpredictability composite appeared to be driven by self-reported perceived unpredictability and not by the interview-based measure which captured changes in youths' living situations. Likewise, the main effect of violence exposure appeared to be driven by neighborhood violence and not involvement or witnessing fights. For working memory updating, interaction effects involving violence exposure were driven by involvement or witnessing fights but not perceptions of neighborhood violence. All components of poverty exposure measures, except self-reported perceptions of resource scarcity, showed interactive effects with task version on working memory updating. Our second set of secondary analyses compared violence to poverty exposure in the same model. Findings suggest, at least when directly comparing interaction terms, violence exposure  $\times$  task version predicted the working memory updating equalization effect whereas the poverty  $\times$  task version interaction did not. Although interesting, all secondary analyses are exploratory, and should be interpreted with caution. Nonetheless, these analyses highlight the need to measure multiple dimensions of adversity, compare them, and consider each measure's report format and source.

Taken together, our equalization patterns of results are striking compared to the backdrop of developmental science, which almost exclusively reports lowered cognitive performance in people exposed to harsh rearing conditions. We also document lowered performance among adversity-exposed youth, but this effect was specific to tests with abstract stimuli, at least for working memory updating. When testing materials have more ecological relevance, adversity-exposed youth perform about as well as youth from affluent backgrounds. These effects have both practical and theoretical implications. Practically, these effects support the idea that ecologically relevant testing materials may "even the playing field" for adversity-exposed youth. That is, pending replication and extension, our data suggest leveraging ecologically relevant materials in education may hold promise for closing achievement gaps<sup>6,7,9</sup>. However, our data also suggest the effect of ecological content may be specific to particular abilities—ecological content actually *lowered* attention shifting performance for everyone compared to abstract content. This negative effect is consistent with other studies reporting lowered performance on tests with ecologically relevant content<sup>9,37,38</sup>.

Theoretically, our findings support the notion that hidden talents may emerge when using ecologically relevant testing materials, but they also raise questions. In particular, why should ecologically relevant stimuli equalize working memory updating performance for adversity-exposed youth compared to youth with reduced exposures? There are at least three possibilities. First, both impairment and enhancement processes could operate in parallel. While abstract working memory updating performance is impaired, ecological working memory updating performance is enhanced. However, this explanation is difficult to reconcile unless performance on tests with different content rely on or recruit different cognitive processes. If this is true,

those mechanisms that process abstract content may be impaired but those that process ecological content are intact or enhanced. Second, there may not be impaired or enhanced processes at all. Instead, ecologically relevant stimuli may activate intact, but normally dormant or underrecruited, working memory updating processes. Thus, when tested under abstract testing conditions, cognitive processes operate at a lower level (remain inactive) but operate normally with ecological testing materials. This makes sense if adversity limits energy and/or cognitive resources. Under such constraints, it may be adaptive to activate expensive cognitive processes only when necessary. Third, ecological testing materials may remove barriers, such as testing anxiety, that disadvantage adversity-exposed youth. Unlike the selective activation of cognitive resources, the additional stressors of testing are removed, allowing cognitive mechanisms to function normally. However, this explanation is difficult to reconcile with experimental work<sup>18,19</sup>, which finds that mild stress/uncertainty improves performance. Future research should use research designs that can disentangle these explanations.

Our study has several strengths and limitations. First, we leveraged a within-person design, which illuminated the effects of ecological testing materials in the same person. However, the correspondence between each version of our tests was modest for working memory updating ( $r = 0.46$ ) and low for attention shifting ( $r = 0.17$ ). Although the task structure was the same, it is unclear to what extent the ecological tests measured the same construct, especially for attention shifting. Second, although we found some effects of ecological testing materials, we measured previously documented hidden talents. To explore generalizability of ecological content across abilities, future research may create ecological versions of tests known to disadvantage adversity-exposed youth, such as inhibition<sup>19</sup> or working memory retrieval<sup>19</sup>, and compare performance to abstract versions. Third, although we used interviews and self-report methods, all of our childhood adversity measures were cross-sectional and some were retrospective. These measurement limitations may be partially offset because our sample was still living in their developmental environment. However, this does not replace the need for prospective, longitudinal measures and cannot eliminate all biases associated with retrospective reporting<sup>39</sup>. Thus, it remains unclear to what extent the current environment, rather than the developmental environment, is driving our findings.

Until recently, developmental science has almost exclusively reported lowered cognitive performance in people from harsh conditions. The hidden talents framework has challenged this notion and inspired researchers to document the strengths and abilities of people from adversity. Though much is yet to be done, our data have exciting practical implications, such as developing ecologically relevant materials for education, that might “even the playing field” for adversity-exposed youth. Theoretically, our data help to refine the hidden talents framework and illuminate future research questions surrounding the unique skills and abilities of people living in harsh conditions.

## Method

### Participants

The current study collected data from 681 adolescents ( $M_{age} = 13.63$ ,  $SD_{age} = 0.83$ ) from two sources across Salt Lake City, Utah. The first was a middle school and the second were 5 after-school clubs. The middle school sample (seventh and eighth graders,  $N = 583$ ; 290 females) formed 85.61% of the total sample and is socioeconomically and ethnically diverse. According to school district records, 40.31% of students were receiving economic assistance based on family income and/or homelessness (i.e., free or reduced-price lunch or fee waivers). The clubs ( $N = 98$ ; 32 females) contained a wider age range ( $M_{age} = 14.22$ ,  $SD_{age} = 1.42$ ). 57.14% were receiving free or reduced-price lunch and were also ethnically diverse.

We aggregated data from both sources and applied a set of five exclusion criteria. We applied these criteria before inspecting or testing any relations between predictors and outcomes. Specifically, we excluded participants: 1) with incomplete cognitive task data, 2) with apparent cognitive impairments (e.g., head injuries, disabilities, drug use), 3) who missed two or more trap questions, 4) exposed to extreme distractions during the assessment, and 5) who received more than 60 minutes per day of special education services. After exclusions, the final sample contained 618 adolescents ( $M_{age} = 13.62$ ,  $SD_{age} = 0.81$ ; 298 females; 87.38% from the middle school).

### Procedure

Data collection involved six parts: consent/assent, demographics form, cognitive testing, structured interview, questionnaire, and a debriefing. Consent was obtained from the primary parent or guardian before assessments and assent was obtained from the participant at the beginning of the session. On the demographics form, participants indicated if they were left-handed, their birthdate, and their racial and ethnic background.

Cognitive testing included abstract (i.e., standard) and ecological versions of an attention-shifting and working memory updating task. Task order (shifting versus updating) and version (abstract versus ecological) were counterbalanced across participants. Some participants started with the attention shifting and then moved to the updating tasks. For those starting with the shifting (or updating) tasks, some started with the abstract version and then completed the ecological version (and vice versa). Regardless of which task was first, participants always completed both versions of one task before moving to the other task.

The interview contained two parts. The first part was about participants' home and family life, such as their family composition, the number of adults in the household, and occupation and education information

about caregivers. Participants also reported on the number of romantic partners their caregivers had and if these partners lived with the participant. The second part was about residential mobility. Participants listed all the homes they lived in from birth to age 13. For each home, they reported at which ages they lived in the home. If they moved, they also reported any changes in their family structure (loss or gain of adults in the home), social network (e.g., a new school), and/or town or city.

The questionnaire contained measures focused on the childhood environment. The measures included perceived unpredictability in the home, exposure to neighborhood violence, access to material resources, direct exposure to violence and fighting, and a child version of socially desirable responding.

Procedures across the middle school and after-school clubs were the same, with a few minor differences. First, participants in the clubs completed the tasks before the interview, but the opposite occurred in the middle school. The second was that, in the middle school, the questionnaire was completed in a separate session. Whereas the tasks and interview were completed in a small room, alone with the interviewer, the questionnaire was completed approximately two weeks later in a mass testing context (large computer room). This was done because the middle school was on a strict class schedule, which limited the length of individual sessions. In the clubs, sessions were conducted in the summer during flexible hours.

## Measures

### *Attention Shifting*

We measured attention shifting with two versions of the color–shape task<sup>18,40,41</sup>, which involves categorizing stimuli according to one of two rules (see below). On each trial, the rule is displayed on the top of the screen. Using arrow keys, participants categorized the stimulus as quickly and accurately as possible. Response options appeared at the bottom of the screen (see Figure 1).

On the abstract (standard) attention-shifting task, participants categorized colored (blue or yellow) shapes (squares or triangles) according to the ‘color’ or ‘shape’ rule (see Figure 1a). For example, if a blue triangle appeared on the screen and the rule was “color”, participants categorized the blue triangle as blue. However, if the rule was “shape”, participants instead categorized the blue triangle as a triangle.

The ecological attention-shifting task replaced shapes and colors with faces (see Figure 1b). Faces were either male or female (gender), and were happy or angry (emotion). If a happy female face appeared on screen and the rule was ‘gender’, for example, participants categorized the face as female. If the rule was ‘emotion’, participants instead categorized the face as happy. The stimuli for the ecological attention-shifting task were selected from the NimStim library<sup>42</sup>. The same female and male models were used for happy and

angry expressions.

For both versions, there were two types of trials: repeat and switch. Repeat trials use the same rule as the previous trial. If the current and previous rule is ‘color’, for instance, the current trial is a repeat trial. Switch trials are trials where the rule changes from the previous trial. If the current rule is ‘color’ but the previous was ‘shape’ (or vice versa), the current trial is a switch trial.

Accuracy on attention-shifting tasks is normally high, and was in the current sample for both versions ( $M_{standard} = 89.88\%$ ,  $SD_{standard} = 8.44\%$ ;  $M_{ecological} = 86.53\%$ ,  $SD_{ecological} = 12.25\%$ ). Thus, instead of accuracy, attention-shifting ability is measured by comparing reaction times on repeat trials to switch trials. Reaction times on repeat trials are normally faster than switch trials; average repeat trial reaction times on the abstract and ecological versions were 1195.29 ms ( $SD = 375.59$  ms) and 1279.53 ms ( $SD = 461.01$  ms), respectively. On switch trials, average reaction times were 1309.23 ms ( $SD = 446.11$  ms) and 1589.07 ms ( $SD = 589.15$  ms) on the standard and ecological versions.

To calculate attention-shifting scores, we first cleaned trial-level data.<sup>41,43</sup> We removed fast and slow reaction times and removed trials after incorrect responses (see [attention shifting data cleaning procedure](#) for details). We then subtracted the average reaction time on valid repeat trials from the average reaction time on valid switch trials to calculate switch cost scores ( $M_{switch} - M_{repeat}$ ). The switch cost measures the relative delay in reaction time caused by switching between categorization rules. Average switch costs on the standard and ecological versions were 134.07 ms ( $SD = 203.66$  ms) and 287.53 ms ( $SD = 255.46$  ms), respectively ( $r = 0.17$ ,  $n = 618$ ). Higher switch costs mean longer reaction times on switch trials compared to repeat trials. Larger switch costs indicate less efficient attention shifting.

### **Working Memory Updating**

Working memory updating was measured using two versions of the continuous counters task<sup>19,44,45</sup>, in which participants track a sequence of stimuli containing three different objects (see Figure 1). Participants keep running counts of each object in sequences nine items long. Each object was displayed for one second at a time. At the end of each sequence, participants reported the number of times each object appeared. For each round, there were three possible correct answers, one for each object type.

In the abstract (standard) version, participants kept running counts of squares, circles, and triangles (see Figure 1c). The participant must keep track of each shape, updating counts as new shapes appeared, and report them at the end of the sequence. The ecological version followed the same format, but participants kept running counts of three different real-world objects (see Figure 1d): one positive (five-dollar bill), one negative (angry face), and one neutral (Salt Lake City bus).

Participants completed five rounds of both versions. The maximum score was 15 for each version (five rounds  $\times$  three possible correct). On the abstract and ecological versions, participants averaged 10.5 ( $SD = 2.7$ ) and 10.95 ( $SD = 2.55$ ), respectively ( $r = 0.46$ ). Higher scores indicate better working memory updating performance.

### *Environmental Unpredictability*

Our measure of environmental unpredictability integrated information from the interview and questionnaire. The interview contained three variables of interest. The first measured the level of exposure to living in a biologically non-intact family from birth to age 13. Interviewers asked participants about their family composition and the adults that lived with them in the home. If a participant's family composition changed, for example, because of a divorce, interviewers recorded the ages when participants experienced the change and for how long it persisted. A family was considered non-intact if the participant did not live with both biological parents during any period from birth to age 13. Based on the earliest age of disruption, we indexed the onset of exposure to a non-intact family by creating a four-level variable from 0 to 3. A 0 indicated the participant lived in an intact family (lived with both biological parents) from birth until age 13 ( $n = 336$ ); a 1 indicated that a disruption occurred after the age of 5 ( $n = 89$ ); a 2 indicated that a disruption occurred before age 5 ( $n = 132$ ); and a 3 indicated the participant was born into a non-intact family ( $n = 56$ ). Five participants had missing information.

The second interview variable tabulated the number of different parental figures (defined as spouses or boyfriends/girlfriends of the participant's mother and father) who lived in the same home with the participant. Trained interviewers tabulated the total number of co-residing parental figures, excluding the participant's biological mother and father ( $M = 1.36$ ,  $SD = 1.8$ ).

The third interview variable focused on residential changes from birth to age 13 (reported retrospectively). Specifically, interviewers probed participants about the number of different homes they lived in and, in particular, whether moving to a new home involved living with new or different adults. On average, participants experienced 1.35 ( $SD = 1.66$ ) that involved changes in residential adults.

The questionnaire included a measure of perceived childhood unpredictability<sup>18,19,46</sup>. Participants responded to six items that assessed exposure to unpredictability up to age 13: 1) "Things were pretty calm and stable in my house" (reversed); 2) "People moved in and out of my house a lot"; 3) "I had a hard time knowing what my parent(s) or other people in my house were going to say or do from day-to-day"; 4) "My parent(s) argued or fought a lot with each other or with other people"; 5) "My parent(s) changed jobs a lot"; 6) "My parental situation changed a lot (for example, divorce or separation of parents, parents starting new

romantic relationships, parents leaving the home)". Participants rated each item on a scale from 1 (never true) to 5 (very often true). Responses to these items were averaged to create a self-reported perceived childhood unpredictability measure ( $M = 1.99$ ,  $SD = 0.73$ ,  $\alpha = 0.72$ ). Higher scores indicated more perceived exposure to unpredictable environments.

To create a composite measure of unpredictability, we aggregated across the interview and questionnaire. First, we created an interview-based measure of unpredictability using family disruptions, exposure to different parental figures, and residential moves involving new adults living in the home. We standardized these interview measures and then averaged them together ( $\alpha = 0.77$ ). We then aggregated this interview measure with self-reported perceived unpredictability from the questionnaire, standardizing each before averaging ( $r = 0.5$ ). Higher scores indicated more exposure to unpredictability.

### *Violence Exposure*

We measured violence exposure using the Neighborhood Violence Scale (NVS)<sup>26,47</sup> and two items about witnessing and involvement in physical fights. The NVS contains 7 items rated on a scale from 1 (never true) to 5 (very often true): 1) "I grew up in a safe neighborhood"; 2) "Crime was common in the neighborhood where I grew up"; 3) "In the neighborhood where I grew up, people had plenty of money" (reversed); 4) "In the neighborhood where I grew up, physical fights were common"; 5) "In the neighborhood where I grew up, shootings or stabbings occurred"; 6) "In the neighborhood where I grew up, most people felt unsafe walking alone after dark"; 7) "Where I grew up, it was important to be able to defend yourself against physical harm". The physical fighting items asked about witnessing fights: "Based on your experiences at school and in your neighborhood, how many times did you see or hear someone being beaten up in real life?"; and involvement in fights: "How many times were you in a physical fight at school or in your neighborhood?". Both items were rated on a scale from 1 (0 times) to 8 (12 or more times).

Responses on the NVS were averaged to create a composite measure ( $M = 1.94$ ,  $SD = 0.75$ ,  $\alpha = 0.85$ ). Likewise, the two fighting were averaged together ( $M = 2.31$ ,  $SD = 1.43$ ). To create a violence exposure composite, the NVS composite and fighting composite were standardized and then averaged ( $r = 0.6$ ). Higher scores indicated more exposure to violence.

### *Poverty Exposure*

We assessed each participant's poverty exposure using three information sources: interview-assessed parent education and occupational prestige, school or club-provided economic information, and perceived resource scarcity.

During the interview, participants reported their parents' education and primary occupation. For parental education level, participants reported whether their mother and father graduated high school, went to college, and if they obtained a bachelor's degree or higher. Responses to each question were coded as 0 = 'no' and 1 = 'yes'. These items were summed to create a four-level variable for each parent where 0 = did not graduate high school, 1 = graduated high school but did not attend any college, 2 = graduated high school and completed some college, and 3 = went to college and obtained a bachelor's degree or higher. Mother and father education were averaged to create a parental education score ( $M = 1.63$ ,  $SD = 0.91$ ).

Participants also reported their parents' occupation. Participants described, in general terms, their mother's and father's jobs and the kind of work involved. Trained coders cross-referenced these descriptions with 2010 census job codes<sup>48</sup>. We calculated mother and father occupational prestige by matching census codes with prestige ratings from an established occupational prestige rating dataset<sup>49</sup>. Prestige codes ranged from 0-100, with higher scores indicating higher occupational prestige for a particular job code. Mother and father prestige codes were averaged to create a parent occupational prestige score ( $M = 41.35$ ,  $SD = 10.57$ ).

We also obtained school or club-provided economic information. The middle school provided four information sources: 1) whether or not students were receiving free lunch ( $n = 159$ ); 2) whether or not they were receiving reduced-price lunch ( $n = 45$ ); 3) whether they were receiving fee waivers ( $n = 103$ ); and 4) if they were experiencing homelessness ( $n = 32$ ). These information sources were compiled to create a binary economic disadvantage variable. A participant was economically disadvantaged if they were receiving any benefits or experiencing homelessness ( $n = 215$ ). In the afterschool clubs, free or reduced-price lunch status was provided by self-report, staff, and club enrollment forms and cross-referenced to determine economic disadvantage in the clubs ( $n = 41$ ). Overall, 256 participants were receiving economic assistance, 318 were not, and 44 had missing economic information.

Participants also reported their perceived level of resource scarcity. They rated seven items from 1 (never true) to 5 (very often true): 1) "Your family had enough money to afford the kind of home you all needed"; "Your family had enough money to afford the kind of clothing you all needed"; 3) "Your family had enough money to afford the kind of food that you all needed"; 4) "Your family had enough money to afford the kind of medical care that you all needed"; 5) "I felt well-off (rich, wealthy) compared to other kids in my school"; 6) "I felt well-off (rich, wealthy) compared to other kids in my neighborhood"; and 7) "Your family struggled to make ends meet (get by financially)". The first six items were reversed, and all items were averaged to create a perceived resource scarcity measure ( $M = 2.2$ ,  $SD = 0.67$ ,  $\alpha = 0.83$ ). Higher scores indicated more perceived resource scarcity.

To create a poverty exposure composite, we combined parent education and occupational prestige (re-



versed), school/club reported economic disadvantage, and perceived resource scarcity. First, we standardized and averaged parent education and occupation ( $r = 0.47$ ). Then, we recoded school-provided economic codes to be equal to half of a standard deviation above and below the mean (in terms of z-scores). Specifically, those receiving free or reduced-price lunch were coded as .5, and those who were not were coded as -.5. To create a final composite, we averaged parent education/occupation (standardized), perceived resource scarcity (standardized), and the recoded school/club economic disadvantage. Higher values indicated more exposure to poverty.

### ***Control Variables***

We included two control variables: age and interviewer-rated distractions during testing. We included these because older adolescents tend to score higher on cognitive tests than younger adolescents<sup>50–52</sup>, and disruptions during the tasks could negatively impact performance. Interviewers completed quantitative ratings of the testing environment including 1) the level of noise; 2) the level of chaos; 3) how often there were interruptions; and 4) how often the participant appeared distracted. Interviewers rated each item on a scale from 1 (not at all) to 7 (much more than average). Items were averaged to create a composite ( $M = 1.3$ ,  $SD = 0.46$ ), where higher scores indicate more distractions.

We also included the Child Social Desirability (CSD) scale<sup>53</sup>. The measure includes 14 items such as “Have you ever broken a rule?” and “Do you sometimes feel angry when you don’t get your way?”. Participants indicated 1 (yes) or 0 (no) for each item. More socially desirable responses were scored as 1. Items were then summed to create a social desirability index ( $M = 5.36$ ,  $SD = 3.1$ ,  $\alpha = 0.77$ ).

### **Multiverse Setup**

We used a multiverse analysis approach for all analyses<sup>29,30,33,34</sup>. Researchers face many data processing decisions. These decisions fall into two classes: arbitrary and non-arbitrary decisions. Non-arbitrary decisions have clear theoretical, statistical, or practical justifications and are superior to alternatives. Arbitrary decisions are those that have (sometimes many) equally defensible alternatives. For example, excluding outliers at  $\pm 2.5$  or 3 standard deviations above (or below) the sample mean could be equally defensible. Instead of cherry-picking one or getting pigeonholed with the other, multiverse analysis allows researchers to explicitly generate all possible combinations of (arbitrary) data processing decisions and analyze them together. In turn, researchers can systematically evaluate the robustness or sensitivity of their analysis across all possible data processing decisions.

### Arbitrary Decisions

We identified six arbitrary data processing decisions. First, we conducted analyses with and without the after-school club sample. We did this because our procedure was slightly different across the two samples and the testing environment was more disruptive in the clubs compared with the middle school. Second, we conducted analyses with and without participants with minor distractions during testing. Third, we analyzed the data with and without participants who missed one trap question (those who missed more than 1 were already excluded). Fourth, we analyzed data with and without participants having any amount of special education (participants receiving more > 60 minutes were already excluded). Fifth, we analyzed data with and without participants who performed at or below chance levels of accuracy on the attention-shifting task (see [attention shifting data cleaning procedure](#) for more information).

Our last arbitrary data processing decision pertained to socially desirable responses. It is well-known that some youth tend to describe themselves in socially desirable ways. This raises concerns about the validity of self-reported childhood adversity measures: youth scoring higher on social desirability might underreport or underrate their exposures to negative experiences. Indeed, scores on the CSD were correlated with exposures to unpredictability ( $r = -0.2$ ) and violence ( $r = -0.27$ ), though not with poverty. One way to handle this issue is to statistically remove overlapping variance between social desirability and each predictor, using the leftover variance (e.g., the residual variance in childhood adversity exposures not associated with social desirability) as a predictor. Because it is unclear how social desirability might affect the interpretation of our models, we analyzed our data with residualized and non-residualized predictors.

### Statistical Inferences

Multiverse analysis produces a distribution of effect sizes and test statistics across all combinations of analytic decisions (see Figures 2 and 3). However, although these distributions provide rich descriptive information, they do not lend themselves to straightforward null-hypothesis tests. This is because each individual analysis in a multiverse is not independent. To address this, we used a bootstrapping technique designed for multiverse analysis<sup>30,33,34</sup>. This technique makes it possible to know if a multiverse analysis produced a median effect size larger than expected, given the null hypothesis (the median effect is 0).

For the current study, we calculated overall  $p$ -values for the main effect of adversity and the interaction effects (task version  $\times$  adversity) of our primary analyses. To do so, we created ‘null’ attention-shifting and working memory updating scores<sup>30</sup>. That is, we created outcome scores with the influence of the effect of interest (i.e., main effect or interaction) removed. More specifically, we multiplied the coefficient

of interest (regression coefficient of the main effect of adversity or interaction effect) by each participant's raw predictor score (either adversity scores or task version  $\times$  adversity scores) and subtracted the result from each participant's attention-shifting and working memory scores. Next, we randomly sampled participants (with replacement) to create 500 bootstrapped samples. Then, we ran the same multiverse analysis on each bootstrap sample using the null attention-shifting and working memory scores as outcome variables. Upon completion, we compared the median effects for each bootstrapped sample (500 median effect sizes) with our original median effects. To compute overall  $p$ -values, we divided the number of effect sizes obtained from the bootstrapped multiverse analyses that were larger than our original median effect sizes by the total number of bootstrapped samples. Importantly, we ran this procedure separately for each measure of adversity (unpredictability, violence exposure, and poverty exposure), outcome (attention shifting and working memory updating), and effect of interest (main effect or interaction effect) combination.

## References

1. Ursache, A. & Noble, K. G. Neurocognitive development in socioeconomic context: Multiple mechanisms and implications for measuring socioeconomic status. *Psychophysiology* **53**, 71–82 (2016).
2. Farah, M. J. *et al.* Childhood poverty: Specific associations with neurocognitive development. *Brain Research* **1110**, 166–174 (2006).
3. Noble, K. G., McCandliss, B. D. & Farah, M. J. Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science* **10**, 464–480 (2007).
4. Hackman, D. A., Farah, M. J. & Meaney, M. J. Socioeconomic status and the brain: Mechanistic insights from human and animal research. *Nature Reviews Neuroscience* **11**, 651–659 (2010).
5. Blair, C. & Raver, C. C. Child development in the context of adversity: Experiential canalization of brain and behavior. *American Psychologist* **67**, 309–318 (2012).
6. Ellis, B. J., Bianchi, J., Griskevicius, V. & Frankenhuis, W. E. Beyond risk and protective factors: An adaptation-based approach to resilience. *Perspectives on Psychological Science* **12**, 561–587 (2017).
7. Ellis, B. J. *et al.* Hidden talents in harsh environments. *Development and Psychopathology* 1–19 (2020) doi:[10.1017/S0954579420000887](https://doi.org/10.1017/S0954579420000887).
8. Frankenhuis, W. E. & de Weerth, C. Does early-life exposure to stress shape or impair cognition? *Current Directions in Psychological Science* **22**, 407–412 (2013).
9. Frankenhuis, W. E., Young, E. S. & Ellis, B. J. The hidden talents approach: Theoretical and methodological challenges. *Trends in Cognitive Sciences* **24**, 569–581 (2020).
10. McLaughlin, K. A., Weissman, D. & Bitrán, D. Childhood adversity and neural development: A systematic review. *Annual Review of Developmental Psychology* **1**, 277–312 (2019).
11. Sheridan, M. A. & McLaughlin, K. A. Dimensions of early experience and neural development: deprivation and threat. *Trends in Cognitive Sciences* **18**, 580–585 (2014).
12. Gibb, B. E., Schofield, C. A. & Coles, M. E. Reported history of childhood abuse and young adults' information-processing biases for facial displays of emotion. *Child Maltreatment* **14**, 148–156 (2009).
13. Pollak, S. D. Mechanisms linking early experience and the emergence of emotions: Illustrations from the study of maltreated children. *Current Directions in Psychological Science* **17**, 370–375 (2008).

14. Pollak, S. D., Messner, M., Kistler, D. J. & Cohn, J. F. Development of perceptual expertise in emotion recognition. *Cognition* **110**, 242–247 (2009).
15. Piff, P. K., Kraus, M. W. & Keltner, D. Unpacking the inequality paradox: The psychological roots of inequality and social class. in (ed. Olson, J. M.) vol. 57 53–124 (Academic Press, 2018).
16. Ein-Dor, T. & Perry, A. Full house of fears: Evidence that people high in attachment anxiety are more accurate in detecting deceit. *Journal of Personality* **82**, 83–92 (2014).
17. Shoda, T. M. & McConnell, A. R. Interpersonal sensitivity and self-knowledge: Those chronic for trustworthiness are more accurate at detecting it in others. *Journal of Experimental Social Psychology* **49**, 440–443 (2013).
18. Mittal, C., Griskevicius, V., Simpson, J. A., Sung, S. & Young, E. S. Cognitive adaptations to stressful environments: When childhood adversity enhances adult executive function. *Journal of Personality and Social Psychology* **109**, 604–621 (2015).
19. Young, E. S., Griskevicius, V., Simpson, J. A., Waters, T. E. A. & Mittal, C. Can an unpredictable childhood environment enhance working memory? Testing the sensitized-specialization hypothesis. *Journal of Personality and Social Psychology* **114**, 891–908 (2018).
20. Nweze, T., Nwoke, M. B., Nwifo, J. I., Aniekwu, R. I. & Lange, F. Working for the future: parentally deprived Nigerian children have enhanced working memory ability. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* **62**, 280–288 (2021).
21. Banerjee, A. V., Bhattacharjee, S., Chattopadhyay, R. & Ganimian, A. J. The untapped math skills of working children in india: Evidence, possible explanations, and implications. *Unpublished manuscript* (2017).
22. Greenfield, P. M. & Cocking, R. R. *Cross-cultural roots of minority child development*. (Psychology Press, 2014).
23. Ogbu, J. Origins of human competence: A cultural-ecological perspective. *Child Development* **52**, 413–429 (1981).
24. Richardson, K. Reasoning with raven - in and out of context. *British Journal of Educational Psychology* **61**, 129–138 (1991).
25. Schliemann, A. D. & Carraher, D. W. The evolution of mathematical reasoning: Everyday versus idealized understandings. *Developmental Review* **22**, 242–266 (2002).

26. Frankenhuis, W. E., de Vries, S. A., Bianchi, J. & Ellis, B. J. Hidden talents in harsh conditions? A preregistered study of memory and reasoning about social dominance. *Developmental Science* **23**, e12835 (2020).
27. Rifkin-Graboi, A. *et al.* Caregiving adversity during infancy and preschool cognitive function: Adaptations to context? *Journal of Developmental Origins of Health and Disease* 1–12 (2021) doi:[10.1017/S2040174420001348](https://doi.org/10.1017/S2040174420001348).
28. VanTassel-Baska, J. Achievement unlocked: Effective curriculum interventions with low-income students. *Gifted Child Quarterly* **62**, 68–82 (2018).
29. Steegen, S., Tuerlinckx, F., Gelman, A. & Vanpaemel, W. Increasing transparency through a multi-verse analysis. *Perspectives on Psychological Science* **11**, 702–712 (2016).
30. Simonsohn, U., Simmons, J. P. & Nelson, L. D. Specification curve analysis. *Nature Human Behaviour* **4**, 1208–1214 (2020).
31. R Core Team. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing, 2020).
32. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**, (2015).
33. Orben, A. & Przybylski, A. K. Screens, teens, and psychological well-being: Evidence from three time-use-diary studies. *Psychological Science* **30**, 682–696 (2019).
34. Orben, A., Dienlin, T. & Przybylski, A. K. Social media’s enduring effect on adolescent life satisfaction. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 10226–10228 (2019).
35. Market, R. T. Build trust through criticism.
36. Fields, A. *et al.* Adaptation in the face of adversity: Decrements and enhancements in children’s cognitive control behavior following early caregiving instability. *Developmental Science* doi:[10.1111/desc.13133](https://doi.org/10.1111/desc.13133).
37. Muskens. Hidden obstacles in education for students from low socioeconomic backgrounds: (2019). doi:[10.26481/dis.20191115mm](https://doi.org/10.26481/dis.20191115mm).
38. Duquenois, C. Fictional Money, Real Costs: 64.

39. Reuben, A. *et al.* Lest we forget: Comparing retrospective and prospective assessments of adverse childhood experiences in the prediction of adult health. *Journal of Child Psychology and Psychiatry* **57**, 1103–1112 (2016).
40. Miyake, A. & Friedman, N. P. The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science* **21**, 8–14 (2012).
41. Friedman, N. P. *et al.* Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General* **137**, 201–225 (2008).
42. Tottenham, N. *et al.* The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research* **168**, 242–249 (2009).
43. Miyake, A. *et al.* The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology* **41**, 49–100 (2000).
44. Unsworth, N., Fukuda, K., Awh, E. & Vogel, E. K. Working Memory Delay Activity Predicts Individual Differences in Cognitive Abilities. *Journal of Cognitive Neuroscience* **27**, 853–865 (2015).
45. Unsworth, N. & Engle, R. W. Speed and accuracy of accessing information in working memory: An individual differences investigation of focus switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **34**, 616–630 (2008).
46. Szepeswol, O., Simpson, J. A., Griskevicius, V. & Raby, K. L. The effect of unpredictable early childhood environments on parenting in adulthood. *Journal of Personality and Social Psychology* **109**, 1045–1067 (2015).
47. Frankenhuys, W. E. & Bijlstra, G. Does exposure to hostile environments predict enhanced emotion detection? *Collabra-Psychology* **4**, 18 (2018).
48. Hout, M. Americans’ occupational status reflects the status of both of their parents. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 9527–9532 (2018).
49. Hout, M., Smith, T. W. & Marsden, P. V. Prestige and socioeconomic scores for the 2010 census codes. (2016).
50. Best, J. R. & Miller, P. H. A developmental perspective on executive function. *Child Development* **81**, 1641–1660 (2010).
51. Best, J. R., Miller, P. H. & Naglieri, J. A. Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences* **21**, 327–336 (2011).

- 733 52. Huizinga, M., Dolan, C. V. & van der Molen, M. W. Age-related change in executive function: De-  
734 velopmental trends and a latent variable analysis. *Neuropsychologia* **44**, 2017–2036 (2006).
- 735 53. Miller, P. H. *et al.* Test-retest reliability of a short form of the children’s social desirability scale  
for nutrition and health-related research. *Journal of Nutrition Education and Behavior* **46**, 423–428  
736 (2014).