finding

Can a visual values-affirmation intervention improve test scores of students in areas affected by crisis?

Daniel D. Shephard, Ali Osseiran, & Fadi Makki

Method and Analysis

Expanded Details of Analysis Methodology

The primary analyses were preregistered before the trial was conducted (AEARCTR-0003081; https://doi.org/10.1257/rct.3081-1.0). The analyses assessed the impact of the values-affirmation task on the performance of the intervention group relative to the control group. As such, for each subject, our primary analytical model regressed each participants' (*i*) exit test score (*y*) on a treatment dummy (*t*) while controlling for the baseline placement test score (*a*) and our stratification variable as educational center fixed effects (*b*). The effect of treatment is represented by the coefficient of treatment (β_1). Performance on each academic subject was analyzed separately.

$$y_i = \beta_1 t_i + \beta_2 a_i + \beta_2 b_i + \varepsilon \quad (1)$$

In addition to the primary analyses, we investigated potential moderation effects for the following covariates: being female, being 18 years old or older, and being Syrian. A secondary model analyzed the outcomes adjusted for two binary covariates—being female (*c*), being 18 years old or older (*d*)—and whether the participant reported being of Syrian nationality. The preregistered analysis planned to use age as a continuous variable; however, given the potential for adult students to be stereotyped, we created a binary variable for analysis.

$$y_i = \beta_1 t_i + \beta_2 a_i + \beta_2 b_i + \beta_3 c_i + \beta_4 c_i \, \varepsilon \quad (2)$$

All analyses were conducted using an ordinary least squares model with adjusted standard errors for heteroscedasticity using the HC2 adjustment implemented by the R packages: lmtest v0.9-35 and sandwich v2.4-0.

Expanded Results

The treatment and control groups were balanced based on the covariates and baseline test performance using a Omnibus design-based balance test (see Table S1).

Primary Effects

Participants in the intervention group scored 6.9 percentage points more on their Arabic test compared with participants in the control group (0.069, 95% CI [0.001, 0.137]); this is the equivalent of an effect size of 0.27 standard deviations. This effect remained when adjusting for covariates (0.070, 95% CI [0.001, 0.139]). However, there was no effect in the main model or covariate model for test scores on English (−0.016, 95% CI [−0.080, 0.048]; −0.010, 95% CI [−0.075, 0.054]) or math (−0.001, 95% CI [−0.055, 0.054]; 0.000, 95% CI [−0.056, 0.057]; see Figure 1 in the main text and Table S2).

Moderation Effects

Syrian refugee participants benefited from the intervention by scoring 25.6 percentage points higher on the Arabic test than non-Syrians participants did (0.256, 95% CI [0.001, 0.511]). The moderation effect of age was not significant, with the estimated intervention effect being only 3.5 percentage points higher among adult participants (0.035, 95% CI [−0.108, 0.177]). We find no moderation by gender on the intervention's effect on participants' Arabic test scores (−0.097, 95% CI [−0.240, 0.047]). There was no support for gender moderation of the lack of impact of the intervention on math scores, despite a positive coefficient for the interaction of the treatment and being female (0.017, 95% CI [−0.099, 0.133]). There was also no support for age-based moderation effects on the mathematics outcome. See Figure 4 in the main text and Table S3.

Table S1. Omnibus design-based balance test, $\chi^2(6) = 1.821$, $p = .935$

Independent variable	Control	Treatment	Diff.	Std. Diff.	*Z*	*p*
Adult learner (18 years old or older)	0.482	0.458	−0.024	−0.048	−0.298	.766
Female	0.594	0.533	−0.061	−0.121	−0.754	.451
Syrian	0.896	0.919	0.023	0.080	0.548	.583
Baseline math	0.227	0.234	0.007	0.039	0.262	.793
Baseline English	0.033	0.042	0.009	0.090	0.653	.514
Baseline Arabic	0.364	0.345	−0.019	−0.082	−0.534	.593

Note. The first three variables are presented as proportions of the sample; the last three baseline variables are the proportions of correct answers. The standardized differences are the differences divided by the standard deviation (Cohen's *d*).

Table S2. Values-affirmation treatment effect on Arabic test scores (proportion correct)

					Confidence interval	
Independent variable	Estimate	*SE*	*t*	*p*	2.5%	97.5%
Treatment	0.069*	0.034	2.012	.046	0.001	0.137
Baseline Arabic	0.576***	0.083	6.934	.000	0.412	0.740
Chiyah	−0.377***	0.063	−5.950	.000	−0.502	−0.251
Haret Hreik	−0.179*	0.074	−2.418	.017	−0.325	−0.033
Hay El Sellom	−0.008	0.075	−0.103	.918	−0.155	0.140
Kamed El Laouz	−0.309***	0.078	−3.950	.000	−0.464	−0.154
Mashghara	−0.257***	0.060	−4.317	.000	−0.375	−0.139
Tyre	−0.266***	0.062	−4.313	.000	−0.387	−0.144

Note. Estimates are the proportion of the total possible score on the Arabic test. The intercept is not displayed to ease legibility. Robust standard errors (SEs) are presented. The excluded reference stratum is Bazourieh.

*$p < .05$. ***$p < .001$.

Table S3. Syrian refugee status moderation effect on Arabic test scores (proportion correct)

					Confidence interval	
Independent variable	Estimate	*SE*	*t*	*p*	2.5%	97.5%
Treatment × Syrian	0.256*	0.129	1.988	.049	0.001	0.511
Treatment	−0.162	0.123	−1.317	.190	−0.406	0.081
Syrian	−0.161*	0.070	−2.303	.023	−0.298	−0.023
Baseline Arabic	0.588***	0.082	7.181	.000	0.426	0.750
Chiyah	−0.342***	0.065	−5.241	.000	−0.470	−0.213
Haret Hreik	−0.167*	0.067	−2.510	.013	−0.299	−0.036
Hay El Sellom	0.024	0.078	0.304	.761	−0.130	0.177
Kamed El Laouz	−0.281***	0.080	−3.495	.001	−0.440	−0.122
Mashghara	−0.226***	0.061	−3.688	.000	−0.348	−0.105
Tyre	−0.235***	0.070	−3.371	.001	−0.372	−0.097

Note: Estimates are the proportion of the total possible score on the Arabic test. The intercept is not displayed to ease legibility. Robust standard errors (*SE*s) are presented. The excluded reference stratum is Bazourieh.

*$p < .05$. ***$p < .001$.