Detailed measures descriptions

*Candidate mediators:*

*Attention bias to threat* was captured by the difference in reaction times to neutral versus angry faces displayed by the visual dot-probe task (Amin et al., 2004; Steinberg et al., 2013). Greater values signify shorter response times to angry rather than neutral expressions. Each trial consisted of a pair of facial expressions from the same actor, displayed side-by-side, with a left- or right-pointing arrow appearing in place of one of the faces. The participating child was instructed to press the corresponding arrow key on a keyboard to identify behind which face the arrow flashed. The task consisted of 16 practice trials, then two experimental blocks of 80 trials each, where each trial began with a central fixation point displayed for 500ms, then the pair of faced displayed for 500ms, and finally the arrow displayed for 500ms. On trials where the participant correctly pressed the button, reaction times were averaged and compared by expression type. Faster reaction times to correctly identify the arrow behind angry faces rather than neutral faces signaled greater attention bias to threat.

*Implicit emotion regulation* was captured by several metrics from the emotional Stroop task (Ben-Haim et al., 2016; Egner et al., 2008; Etkin et al., 2006). Each trial consisted of an image with a happy or fearful facial expression displayed with the word “happy” or “fear” overlaid over the image. In congruent trials, the emotional valence of the face matched the label displayed, whereas in incongruent trials, the emotion label was inconsistent with the facial expression, and required the child to correctly identify the valence of the facial expression despite a conflicting label. We used three measures of implicit emotion regulation derived from this task. Using only trials with correctly identified facial expressions, we contrasted reaction times on incongruent versus congruent trials with fearful faces and happy faces separately. We also included a variable for adaptation to emotional conflict. Faster reaction times are expected on incongruent trials that are preceded by incongruent trials rather than congruent trials, since cognitive control over emotional conflict is expected to be elicited. Adaptation to emotional conflict was operationalized as the difference in reaction times on incongruent trials that were preceded by congruent trials versus reaction times on incongruent trials preceded by incongruent trials (Kim et al., 2021). Higher values on this contrast signal greater adaptation to emotional conflict.

*Theory of mind* was measured with a Theory of Mind task, adapted from a previously implemented task to detect differences between cognitive and affective theory of mind dimensions (Heleniak & McLaughlin, 2020; Schlaffke et al., 2015). Cartoons depicting vignettes of cooperation or cooperation to deceive were shown to children who were asked to predict the conclusion of each story and respond to questions about the story. Of the twelve vignettes, four involved two characters cooperating to achieve a common goal, four depicted a story where the two characters deceived each other, and in the remaining four, the two characters together deceived a third character. The vignettes were displayed in experimental blocks with each of three types of stories displayed in random order. Each block began with a set of instructions corresponding to cognitive theory of mind, affective theory of mind, or physical causality conditions. Cognitive theory of mind represents the ability of participants to understand thoughts, beliefs, and intentions of the characters in the cartoon while affective theory of mind gages whether the children can accurately interpret the emotional state of the characters. Average accuracy metrics on cognitive and affective theory of mind trials were recorded.

*Fear learning* was measured by the difference in the amplitude of skin conductance response (SCR) during the early acquisition phase of a block design fear conditioning and extinction task that has been adapted to the early adolescent population (Shechner et al., 2015). SCR was captured by electrodermal activity and calculated using standard procedures as the difference in the 1–5 s following stimulus onset, with a minimum response of 0.05 microsiemens (μs) (Braithwaite et al., 2013). A blue square and orange diamond were conditioned stimuli. In the initial sequence, the first 10 trials presented the blue square with no aversive reinforcement – the CS- block. Next, 10 trials presented the orange diamond and 8 of the 10 were paired with an aversive sound – the US block. Next, 10 trials showed the orange diamond without the aversive stimulus – the CS+ block. After the initial sequence (of which blocks 2 and 3 were considered pre-acquisition), 9 stimulus blocks (3 reinforced US, 3 non-reinforced CS+ blocks, and 3 CS- blocks) were shown in random order in sets of 3. The difference in SCR on CS+ trials versus CS- trials in the first set of 3 blocks after the pre-acquisition set, corrected for pre-acquisition conductance, was used to capture fear learning.

*Pubertal timing* was assessed using the Tanner staging method (Marshall & Tanner, 1969, 1970). Children were shown sex-specific line drawings conveying stages of development of sexual characteristics (breasts for girls, testes/scrotum/penis for boys, and pubic hair for both). Pubertal timing was the average of the two sex-specific ratings.

*Inhibitory control*, an executive functioning ability to suppress a prepotent response to achieve a longer-term goal, was measured using two tasks. NEPSY Circles & Squares task tested the children’s reaction time on “inhibit” and “switch” tasks (Brooks et al., 2009). The participants were shown a series of circle and square shapes, with some of them shaded in, and asked to read through the shapes to establish a baseline reaction time. In the “inhibit” trials, the participants were asked to say the opposite of the shape presented, regardless of whether it was shaded in or blank. In the “switch” trial sequence, they are asked to say the opposite of the shape if it was blank, and the true shape if it was shaded in. Greater differences in reaction times on “inhibit” and “switch” trial sequences relative to baseline indicate poorer inhibitory control. Stroop task measured the ability of the participating children to accurately read words for colors, even if the color with which the word is presented didn’t match, with greater accuracy conveying greater inhibitory control (Stroop, 1935).

*Language ability and reasoning ability* were measured using the Wechsler Abbreviated Scale of Intelligence (WASI) task (Wechsler, 1999). Language ability was measured with the t-score on the WASI vocabulary subtest. The vocabulary subtest was designed to measure word knowledge and verbal concept formation. Reasoning ability was measured with the t-score on the WASI matrix reasoning subtest, which gages fluid intelligence, broad visual intelligence, classification and spatial ability, knowledge of part–whole relationships, simultaneous processing, and perceptual organization.

Lastly, *reward sensitivity* was assessed using the Piñata task, a child-friendly version of a monetary incentive task (Helfinstein et al., 2013). Animal-shaped piñatas with stars depicted inside appeared on the screen, and the participating children are asked to “whack” each piñata as quickly as possible once it dropped to the middle of the screen by pressing the spacebar on a keyboard. Each piñata was previewed as having 0, 1, 2, or 4 stars inside before it was dropped. If the piñata was hit before it dropped to the middle of the screen, no stars were earned. The stars were earned if the spacebar was pressed within a response window (between 250ms and 300ms after the piñata dropped to the middle of the screen). Earned stars accumulated in a basket at the bottom of the screen. The participating children were told that they will be awarded $10 if they earn enough starts, but ultimately, all children were awarded the $10. The participants were given 22 trials to practice at the start, during which baseline response times were recorded. The task consisted of 132 trials overall, with 6 runs of 22 trials each. The trials were evenly split by reward level with 33 trials each displaying 0, 1, 2, and 4 stars. The total earned stars and the contrast in average reaction times on no-reward (0-star) versus high-reward (4-star) trials measured reward sensitivity, with greater total stars and a greater reaction time contrast conveying greater reward sensitivity.

*2.2.3. Psychopathology symptoms*:

Internalizing symptoms of depression, anxiety, and post-traumatic stress disorder (PTSD) were measured with total scores on child-reported Children’s Depression Inventory-2 (CDI), Screen for Child Anxiety Related Emotional Disorders (SCARED), and UCLA PTSD Reaction Index, respectively (Birmaher et al., 1997; Kovacs, 2011; Steinberg et al., 2004). Externalizing psychopathology outcomes were constructed using the maximum of child and parent reports on attention problem, rule-breaking, and aggression subscales of the Youth Self-Report (YSR) and the Child Behavior Checklist (CBCL) (Achenbach, 1991; Liu et al., 1997). We focused on latent internalizing and externalizing psychopathology outcomes, constructed using a confirmatory factor analysis performed in MPlus Version 8.1 (Muthén & Muthén, 2017) on deciles of scores for depression, anxiety, PTSD, attention problem, rule-breaking, and aggression. The algorithm for internalizing and externalizing composites has been previously described (D. G. Weissman et al., 2020) and is outlined in Supplementary Materials.

Table S.1: Distributions of key variables in the overall sample

|  |  |  |
| --- | --- | --- |
| Characteristic | Overall | % Missing |
| n | 227 |  |
| Age, years, mean(SD) | 11.47 (0.48) | 0 |
| Female biological sex, n(%) | 110 (48.5) | 0 |
| Chronicity of poverty, early childhood, mean(SD) | 0.92 (1.40) | 7.5 |
| Ever poverty, early childhood, n(%) | 80 (38.1) | 7.5 |
| Maternal depression, early childhood, mean(SD) | 23.94 (7.55) | 0 |
| Threat, mean(SD) | 0.01 (0.77) | 0 |
| Deprivation, mean(SD) | 0.01 (0.71) | 0 |
| AB: Attention bias threat, reaction time difference in ms, mean(SD) | -4.91 (35.02) | 5.3 |
| IER: Adaptation to emotional conflict, reaction time difference in ms, mean(SD) | 8.07 (126.39) | 6.6 |
| IER: Stroop - fear, reaction time difference in ms, mean(SD) | -7.40 (88.52) | 6.6 |
| IER: Stroop - happy, reaction time difference in ms, mean(SD) | -5.31 (85.37) | 6.6 |
| ToM: Accuracy on affective trials, mean(SD) | 0.91 (0.10) | 10.6 |
| ToM: Accuracy on cognitive trials, mean(SD) | 0.79 (0.10) | 10.6 |
| FC: SCR to CS+ vs CS-, μs, mean(SD) | 0.18 (0.19) | 15.4 |
| PT: Tanner stage, mean(SD) | 2.21 (0.85) | 15 |
| LA: Language ability, t-score, mean(SD) | 60.06 (8.98) | 0 |
| RA: Reasoning ability, t-score, mean(SD) | 55.56 (9.27) | 0 |
| IC: Reaction time (ms) on inhibit trials, mean(SD) | 5.02 (4.17) | 18.9 |
| IC: Reaction time (ms) on switch trials, mean(SD) | 26.64 (8.89) | 18.9 |
| IC: Accuracy on Stroop task, mean(SD) | 0.33 (0.09) | 10.6 |
| RS: Reaction time (ms) on no- vs high-reward trials, mean(SD) | -28.98 (56.85) | 7.5 |
| RS: Total stars, mean(SD) | 59.82 (14.47) | 7.5 |
| Internalizing symptoms, mean(SD) | 0.00 (1.19) | 6.6 |
| Externalizing symptoms, mean(SD) | 0.04 (2.14) | 6.6 |

% Missing out of 227 with baseline data

Table S.2: HIMA results with outcome and mediator models mutually adjusted for threat and deprivation.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Adversity Exposure | Mediator | Adversity-Psychopathology a | Adversity-Mediator a | Mediator-Psychopathology b | BH-corrected  p-value |
| Standardized β (95% CI) | Standardized β (95% CI) | Standardized β (95% CI) |
| Internalizing | Threat | PT: Tanner stage | 0.16(0.03,0.29)\* | 0.02(-0.11,0.15) | 0.19(0.06,0.32)\*\* | 0.8144 |
| Threat | RS: Reaction time on no- vs high-reward trials | 0.16(0.03,0.29)\* | -0.18(-0.32,-0.04)\* | -0.15(-0.27,-0.03)\* | 0.0212 |
| Deprivation | PT: Tanner stage | 0.30(0.16,0.44)\*\*\* | -0.13(-0.27,0.01) | 0.19(0.06,0.32)\*\* | 0.1373 |
| Deprivation | RS: Reaction time on no- vs high-reward trials | 0.30(0.16,0.44)\*\*\* | -0.08(-0.23,0.07) | -0.15(-0.27,-0.03)\* | 0.3115 |
| Externalizing | Threat | PT: Tanner stage | 0.27(0.14,0.40)\*\*\* | 0.02(-0.11,0.15) | 0.20(0.07,0.33)\*\* | 0.8721 |
| Threat | RS: Total stars | 0.27(0.14,0.40)\*\*\* | -0.18(-0.32,-0.04)\* | -0.05(-0.17,0.07) | 0.8721 |
| Deprivation | PT: Tanner stage | 0.21(0.07,0.35)\*\* | -0.13(-0.27,0.01) | 0.20(0.07,0.33)\*\* | 0.1373 |
| Deprivation | RS: Total stars | 0.21(0.07,0.35)\*\* | -0.08(-0.23,0.07) | -0.05(-0.17,0.07) | 0.6590 |
| p-value \*\*\*<0.001, \*\*<0.01, \*<0.05 | | |  |  |  |  |
| a Adjusted for the other adversity exposure, age, biological sex, early life poverty chronicity, and maternal depression | | | | | |  |
| b Adjusted for threat, deprivation, age, biological sex, early life poverty chronicity, and maternal depression | | | | | |  |
| Standardized beta coefficients represent the change in the outcome associated with a 1-SD change in the predicting variable | | | | | |  |
| BH=Benjamini-Hochberg, preserving false discovery rate at 0.05 | | | |  |  |  |

Table S.3: HIMA results in the overall sample and by sex.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Adversity Exposure | Mediator | Sample | Adversity-Psychopathology a | Adversity- Mediator a | Mediator-Psychopathology b |
|  | Standardized β (95% CI) | Standardized β (95% CI) | Standardized β (95% CI) |
| Internalizing | Threat | Tanner Stage | All | 0.22(0.09,0.35)\*\* | -0.01(-0.13,0.11) | 0.19(0.06,0.32)\*\* |
|  | Boys | 0.17(0.00,0.34)\* | 0.11(-0.05,0.27) | 0.15(-0.04,0.34) |
|  | Girls | 0.29(0.08,0.5)\*\* | -0.17(-0.37,0.03) | 0.28(0.09,0.47)\*\* |
| Threat | Reward Sensitivity (RT contrast) | All | 0.22(0.09,0.35)\*\* | -0.20(-0.33,-0.07)\*\* | -0.15(-0.27,-0.03)\* |
|  | Boys | 0.17(0.00,0.34)\* | -0.19(-0.35,-0.03)\* | -0.11(-0.29,0.07) |
|  | Girls | 0.29(0.08,0.50)\*\* | -0.20(-0.42,0.02) | -0.17(-0.34,0.00) |
| Deprivation | Tanner Stage | All | 0.34(0.2,0.48)\*\*\* | -0.12(-0.25,0.01) | 0.19(0.06,0.32)\*\* |
|  | Boys | 0.33(0.16,0.50)\*\*\* | -0.06(-0.22,0.10) | 0.15(-0.04,0.34) |
|  | Girls | 0.32(0.08,0.56)\*\* | -0.20(-0.42,0.02) | 0.28(0.09,0.47)\*\* |
| Deprivation | Reward Sensitivity (RT contrast) | All | 0.34(0.2,0.48)\*\*\* | -0.12(-0.26,0.02) | -0.15(-0.27,-0.03)\* |
|  | Boys | 0.33(0.16,0.50)\*\*\* | -0.07(-0.24,0.1) | -0.11(-0.29,0.07) |
|  | Girls | 0.32(0.08,0.56)\*\* | -0.21(-0.46,0.04) | -0.17(-0.34,0.00) |
| Externalizing | Threat | Tanner Stage | All | 0.31(0.18,0.44)\*\*\* | -0.01(-0.13,0.11) | 0.21(0.08,0.34)\*\* |
|  | Boys | 0.24(0.07,0.41)\*\* | 0.11(-0.05,0.27) | 0.23(0.03,0.43)\* |
|  | Girls | 0.41(0.23,0.59)\*\*\* | -0.17(-0.37,0.03) | 0.25(0.08,0.42)\*\* |
| Threat | Reward Sensitivity (Total stars) | All | 0.31(0.18,0.44)\*\*\* | -0.02(-0.15,0.11) | -0.19(-0.31,-0.07)\*\* |
|  | Boys | 0.24(0.07,0.41)\*\* | -0.04(-0.23,0.15) | -0.18(-0.35,-0.01)\* |
|  | Girls | 0.41(0.23,0.59)\*\*\* | 0.00(-0.19,0.19) | -0.22(-0.39,-0.05)\* |
| Deprivation | Tanner Stage | All | 0.28(0.14,0.42)\*\*\* | -0.12(-0.25,0.01) | 0.21(0.08,0.34)\*\* |
|  | Boys | 0.27(0.09,0.45)\*\* | -0.06(-0.22,0.1) | 0.23(0.03,0.43)\* |
|  | Girls | 0.28(0.06,0.50)\* | -0.2(-0.42,0.02) | 0.25(0.08,0.42)\*\* |
| Deprivation | Reward Sensitivity (Total stars) | All | 0.28(0.14,0.42)\*\*\* | -0.04(-0.19,0.11) | -0.19(-0.31,-0.07)\*\* |
|  | Boys | 0.27(0.09,0.45)\*\* | -0.12(-0.31,0.07) | -0.18(-0.35,-0.01)\* |
|  | Girls | 0.28(0.06,0.50)\* | 0.07(-0.15,0.29) | -0.22(-0.39,-0.05)\* |
| p-value \*\*\*<0.001, \*\*<0.01, \*<0.05 | | |  |  |  |  |
| a Adjusted for age, biological sex, early life poverty chronicity, and maternal depression | | | | | |  |
| b Adjusted for threat, deprivation, age, biological sex, early life poverty chronicity, and maternal depression | | | | | | |
| Standardized beta coefficients represent the change in the outcome associated with a 1-SD change in the predicting variable | | | | | | |

Table S.4: Distributions of deprivation and threat experiences by biological sex.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Male | Female | p-value |
| N(%) | 117 (51.5) | 110 (48.5) |  |
| Overall deprivation: mean(sd) | 0.12 (0.73) | -0.11 (0.67) | 0.014 |
| **Cognitive deprivation** - reverse-coded count of cognitive stimulation items on the HOME-SF: mean(sd) | 2.64 (1.86) | 2.43 (1.67) | 0.375 |
| **Emotional deprivation** - standardized composite of the CECA and MNBS emotional neglect measures: mean(sd) | 0.18 (0.88) | -0.18 (0.80) | 0.002 |
| **Physical deprivation** - standardized composite of food insecurity and physical neglect subscales of MNBS and CTQ: mean(sd) | 2.62 (0.93) | 2.47 (0.86) | 0.200 |
| Overall threat: mean(sd) | 0.09 (0.79) | -0.08 (0.74) | 0.086 |
| **Count of distinct types** of violence experienced (physical, sexual, domestic violence, witnessing violent crime, victim of violent crime) : mean(sd) | 0.33 (0.84) | 0.22 (0.61) | 0.242 |
| **Summed frequency** ratings of witnessed and experienced violence on VEX-R: mean(sd) | 5.56 (5.92) | 3.97 (4.23) | 0.026 |
| Sum of physical and sexual abuse **severity** on CTQ: mean(sd) | 10.59 (1.36) | 10.56 (2.76) | 0.914 |