Kaufman Scales

The Kaufman Scale characterizes each category of abuse using a number of indices and descriptions and each category is rated from zero to four. Example excerpts of wording from the instrument indicating the most severe rating of 4 include: Physical abuse: "Injuries from intentionality inflicted burns, fractures, injuries requiring hospitalizations"; Neglect: "Five or more forms of neglect" from a list including poor supervision, not provided routine medical care; Sexual abuse: " Vaginal or anal intercourse / penetration vaginal or anal intercourse / penetration"; Emotional abuse: "Exposure to parental drug/alcohol abuse, and in addition, there is evidence of extreme parental rejection (e.g. child called unworthy of love, openly rejected, parent threatens to send child away and/or leave child"); Domestic violence: "Exposure to partner physical violence, with weapon (e.g. knife or other object)".

Table S1. Documented Maltreatment Experience, Severity, Estimated Age of Onset and Duration (in months) at baseline. Note that n and % are not cumulative because of multi-victimisation (i.e. categories are not exclusive).

	MT n =19	***%***
Physical abuse	2	11%
Neglect	13	68%
Sexual abuse	2	11%
Emotional abuse	18	95%
Domestic violence	10	54%
	Mean	*SD*
Severity	7.26	3.93
Age at onset (months)	43.76	54.8
Duration of abuse (months)	135.41	121.68

Childhood Trauma Questionnaire

All children were also administered the Childhood Trauma Questionnaire (CTQ; Bernstein et al., 2003), a child self-report measure assessing emotional and physical neglect, as well as emotional, physical and sexual abuse, yielding separate scores for each domain as well as a composite average score. At baseline all participants were administered the CTQ assessing abuse and neglect over the participants lifetime whereas at follow-up the instruction was changed to report any maltreatment experience in the period since baseline measurement (i.e. maltreatment experiences in the past two years). Total average CTQ score was significantly different between the groups at baseline ($t(35)=-2.61$, $p=.013$; MT: *Mean*=7.92, *SD*=3.24; Non-MT: *Mean*=5.82, *SD*=1.07), but ceased to differ at follow-up measurement ($t(33)=-1.16$, $p=.26$; MT: *Mean*=6.1, *SD*=.84; Non-MT: *Mean*=5.65, *SD*=1.42). A repeated measures analyses showed that the group by time interaction was significant ($F(1,33)=4.51$, $p=.04$), with only the MT group showing a significant decrease in overall CTQ over time (MT: p=.02 ; Non-MT $p=.48$).

AMT stimulus material

Cue words used at T1:

- **Positive:** Friendly, Happy, Respect, Caring, Sunny, Perfect
- **Negative:** Failure, Disliked, Ugly, Useless, Angry, Lonely

Cue words used at T2:

- **Positive:** Fun, Confident, Lucky, Nature, Capable, Peaceful
- **Negative:** Miserable, Excluded, Pain, Upset, Stupid, Terrible

Behavioural In-Scanner Ratings

Participants rated the vividness and difficulty of the ABMs that were recalled in the scanner after each ABM. Analyses of these ratings indicated that the groups were comparable on these parameters as well as time to recall each ABM in scanner (**Baseline -** *RT* Non-MT: Mean= 2.65, SD= .80; *RT* MT: Mean= 2.84, SD= .62, *p*= .52; *Vividness*: Non-MT Mean= 3.73, SD= .75; MT Mean= 3.74, SD= .75, *p*= .97; Difficulty: Non-MT Mean= 2.20, SD= ;.48 MT Mean= 2.35, SD= .51, *p*= .44; **Follow-up** - *RT* Non-MT: Mean= 2.65, SD= .80; *RT* MT: Mean= 2.84, SD= .62, *p*= .51; *Vividness*: Non-MT Mean= 3.50, SD= 1.64; MT Mean=3.42, SD=1.49, *p*= .88; Difficulty: Non-MT Mean= 1.40, SD= 7.1; MT Mean= 1.73, SD= .80, *p*= .28).

Acquisition parameters

During each run, a total of 181 T2* weighted echo-planar (EPI) volumes were acquired, covering the whole brain with the following acquisition parameters: slice thickness: 2mm; TR: 85ms; TE: 50ms; FOV: 192 mm x 192 mm2; 35 slices per volume, gap between slices: 1mm; flip angle: 90°). A high-resolution, three-dimensional T1- weighted structural scan was acquired with a magnetization prepared rapid gradient echo sequence. Imaging parameters were: 176 slices; slice thickness = 1 mm; gap between slices = 0.5 mm; echo time = 2730 msec; repetition time = 3.57 msec; field of view = 256 mm x 256mm2; matrix size = 256 x 256; voxel size = 1 x 1 x 1 mm resolution.

Processing Pipeline

Flexible factorial for cross-sectional data analyses at baseline and follow-up

Data analyses were conducted using the software package SPM12 (www.fil.ion.ucl.ac.uk/spm/software/spm8) implemented in Matlab 2018. After discarding the first 3 volumes of each run to allow for T1 equilibration effects, data were realigned; initially within each run and then across the two runs to the first image of the first run. Data were normalized into MNI space using deformation fields from T1 scan segmentation at a voxel size of 3x3x3mm. The resulting images were smoothed with a 6mm Gaussian filter and high-pass filtered at 128Hz. Fixed-effects statistics for each individual were calculated by convolving boxcar functions modelling the 4 conditions (Positive ABM, Negative ABM, Object recall and Rating) with a canonical hemodynamic response function (HRF). Additionally, the six motion parameters were added to the model as regressors of no interest. In order to further minimize movement-related artefacts, images corrupted due to head motion greater than 1.5mm were removed and replaced by interpolations of adjacent images (overall less than 7% of each participant’s data). For these participants, we included an additional regressor of no interest to model the interpolated scans. Individual participants’ SPMs containing the parameter estimates of the 3 conditions of interest (Positive ABM, Negative ABM, Object recall) were then entered as fixed-effects factors into a repeated measures mixed-effects

ANOVA containing a ‘subject’ and a ‘repetition’ factor for random effects for cross-sectional and longitudinal group analyses of group X time X condition interactions.

Sandwich Estimator (SwE) Toolbox for longitudinal analyses

In addition, we measured areas that showed linear increases over time in both groups. It has been argued that the traditional method modelling longitudinal data as described above makes restrictive assumptions in assuming a common covariance structure for all the voxels in the brain. A more accurate method as compared to the traditional linear mixed effect model that allowed us to investigate areas that increased over time is the Sandwich Estimator (SwE) method, a fast, non-iterative tool for longitudinal and repeated measures data (Guillaume et al., 2014) that is implemented as a SPM Toolbox in SPM 12. The Sandwich Estimator method first estimates the parameters of interest with a simple Ordinary Least Square model and second estimates variances/covariances with the Sandwich Estimator (SwE) accounting for the within-subject correlation existing in longitudinal data.