Supplementary materials for
A comprehensive multi-level analysis of the Bucharest Early Intervention Project: causal effects on recovery from early severe deprivation

Overview
We present the following information in the Supplementary Material, in the order that it is referenced in the main manuscript:

1. Sample characteristics at baseline.
2. Details about scoring and preprocessing of outcome measures and distributions of raw values for outcome measures at each assessment wave.
3. Details and distributions for age of assessment, age of placement in foster care, and stability of placement in foster care.
4. Analyses of potential patterns in missing data.
5. Additional information regarding results of analyses examining sources of variation in the effects of the intervention on children's outcomes, including statistics for results of tests of three-way interactions.
6. Additional information regarding results of analyses examining associations of age of placement and stability of placement with outcomes among children randomized to foster care, including statistics for results of tests of two-way interactions.
7. Supplementary analyses using an extended random effects structure.
8. Supplementary analyses treating age as a categorical variable.

Method

Sample characteristics

Table S1. Demographics and baseline characteristics.

Variable	Mean (SD) or N (%)		χ^2 or Welch's *t* [95% CI]
	Care as usual (n=68)	Foster care (n=68)	
Female	33 (49)	34 (50)	<0.01 [-0.17, 0.18]
Ethnicity			1.46 [-0.06, 0.30]
Romanian	34 (50)	42 (62)	
Rroma	21 (31)	17 (26)	
Other ethnicity	1 (1)	1 (1)	
Unknown	12 (18)	7 (10)	
Age (months)	21.01 (7.43)	20.48 (7.42)	0.41 [-1.99, 3.05]
Birth weight (grams)	2830.33 (567.59)	2718.55 (628.51)	1.03 [-102.76, 326.33]
Gestational age (weeks)	37.57 (1.50)	36.98 (2.62)	1.49 [-0.19, 1.37]
% of lifetime in institutional care	85.96 (21.60)	86.08 (21.69)	-0.03 [-7.46, 7.23]

Outcomes
We present distributions of raw values for each measure in Figures S1–S10.

IQ
We list measures and scoring for IQ in Table S2. As previously described (1, 2), at the 30- and 42-month assessment waves, children completed the Bayley Scales of Infant Development II (3) to assess their cognitive functioning. The Mental Development Index (MDI), a scaled score, ranged from 50 to 150. Children who obtained scaled scores <50 were assigned a numeric MDI score of 49. Raw scores were then assigned an extrapolated age equivalent score to facilitate analyses when scaled scores <50 were obtained (4). Finally, Developmental Quotients (DQs) were computed for each child (i.e., [age equivalent score/chronological age] × 100).

Table S2. Measures of IQ at each wave

Wave	Measure	Scoring
30 months 42 months	BSID-II	Developmental Quotient (DQ)
54 months	WPPSI-R	Full scale Intelligence Quotient (IQ)
8 years 12 years 18 years	WISC-IV	Full scale IQ

Notes. BSID-II = Bayley Scales of Infant Development (3); WPPSI-R = Wechsler Preschool Primary Scale of Intelligence (19); WISC-IV = Wechsler Intelligence Scale for Children–Fourth Edition (20)

EEG data acquisition, processing, and analysis

At the 30-month, 42-month, and 8-year assessment waves (5, 6), EEG was recorded from 12 scalp sites (F3, F4, Fz, C3, C4, P3, P4, Pz, O1, O2, T7, and T8) plus the left and right mastoids using a lycra Electro-Cap (Electro-Cap International Inc., Eaton, OH) with tin sewn-in electrodes, sampled at 512 Hz. EEG was collected in reference to Cz, with AFz serving as the ground. One channel of vertical electrooculogram (EOG) was recorded with tin electrodes placed above and below the left eye to record blinks and eye movements. The EEG and EOG signals were amplified using custom bioelectric amplifiers from SA Instrumentation Company (San Diego, CA). During online recording, a band-pass filter of 0.1 Hz to 100 Hz was applied to all channels. Impedances were kept at or below 10 kΩ. At the 12- and 16-year assessments (7, 8), EEG was recorded using a 64-channel HydroCel Geodesic Sensor Net and a NetAmps 300 amplifier (Electrical Geodesics, Inc., Eugene, OR), sampled at 500 Hz. Impedances of the 64 channels were kept at or below 20 kΩ.

At the 30- and 42-month assessments, EEG was recorded as children observed an experimenter place a number of brightly colored balls in a bingo wheel and spin the wheel for a series of nine trials, each lasting 10 seconds (5). Trials were separated by 10-second intervals in which the experimenter stopped spinning the wheel and changed the number of balls in the wheel to maintain the child's attention. The EEG signal was recorded for the entire 3-minute period, but only data from epochs in which the wheel was being spun were analyzed. At the 8-, 12- and 16- year assessments, EEG was recorded as children sat quietly in a chair, alternating 1-minute epochs of eyes open and eyes closed for a total of 6 minutes. During the eyes-open condition, children were instructed to fixate on a small white cross at the center of a computer screen. Consistent with previous reports from the BEIP, analyses focused on the eyes-open condition (5, 8–10)

All EEG data were processed with the Maryland Analysis of Developmental EEG (MADE) pipeline (11) d(v1.1). The MADE pipeline is implemented in MATLAB and uses the toolbox EEGLAB (12) and its plugins "firfilt" (13), FASTER (14), ADJUST (15), and Adjusted-ADJUST (16). Epochs with voltage values exceeding ±150 μV were rejected. Due to the lower number of channels at the 30-month, 42-month, and 8-year assessments, no interpolation of bad channels was performed and data were processed with the "mini-MADE" processing option, which was designed for low-density recordings and skips steps requiring a larger number of channels. Specifically, the mini-MADE option skips the use of FASTER to identify bad channels and the use of ICA to identify components containing artifact. The 12- and 16-year EEG data were subjected to all MADE pipeline steps, including interpolation of bad channels. Participant data from the eyes-open/eyes-closed paradigm were included if there were at least 10 artifact-free eyes-open trials.

Relative alpha power was computed as follows. First, for each electrode, a fast Fourier transform with a 2-s Hanning window was applied to epoched data in 0.5-Hz bins. Next, the resulting spectral power values (in μV^2) were averaged across trials, and absolute power was computed by taking the natural logarithm of spectral power in each frequency bin. Absolute power reflects the total energy of the signal measured by the scalp electrodes at each frequency (17). In contrast to absolute power, relative power reflects how much a particular frequency band accounts for the total power at a particular scalp electrode (18). Relative power was computed as the proportion of absolute power in the alpha band relative to the total power across 1-45 Hz. The alpha band was defined in line with past reports from the BEIP (age 30-42 months: 6–10 Hz [5]; age 8 years: 7–12 Hz [6]; age 12–16 years: 8–13 Hz [7, 8]). Relative power values were averaged across electrodes F3, F4, C3,

C4, P3, P4, T7, T8, O1, and O2 (i.e., the electrode sites that were common across assessment waves) or their Geodesic Sensor Net equivalents (E12, E60, E20, E50, E28, E42, E24, E52, E35, and E39).

Psychopathology

We list measures and scoring for each type of psychopathology in Table S3.

Table S3. Measures of psychopathology at each wave

Wave	Domain	Measure (informant)	Scoring
30 and 42 months	Disinhibited social engagement disorder	DAI (C)	Sum of disinhibited items (6–8)
	Reactive attachment disorder	DAI (C)	Sum of inhibited items (1–5)
	ADHD	ITSEA (C)	Mean *of activity/inattention* subscale
	Internalizing	ITSEA (C)	Mean of *depression/withdrawal, general anxiety, separation distress,* and *inhibition to novelty* subscales
	Externalizing	ITSEA (C)	Mean *of aggression/defiance* subscale
54 months	Disinhibited social engagement disorder	DAI (C)	Sum of disinhibited items (6–8)
	Reactive attachment disorder	DAI (C)	Sum of inhibited items (1–5)
	ADHD	PAPA (C)	Sum of ADHD items
	Internalizing	PAPA (C)	Sum of anxiety and depression items
	Externalizing	PAPA (C)	Sum of oppositional defiant disorder and conduct disorder items
8 years	Disinhibited social engagement disorder	DAI (C)	Sum of disinhibited items (7, 8, 11a)
	Reactive attachment disorder	DAI (C)	Sum of inhibited items (1, 3a, 4–6)
	ADHD	HBQ (T)	Mean of *inattention* and *impulsivity* subscales
	Internalizing	HBQ (T)	Mean of *overanxious* and *depression* subscales
	Externalizing	HBQ (T)	Mean of *oppositional defiant*, *conduct problems*, *overt aggression*, and *relational aggression* subscales
12 and 16 years	Disinhibited social engagement disorder	DAI (C)	Sum of disinhibited items (7–9)
	Reactive attachment disorder	DAI (C)	Sum of inhibited items (1, 3a, 4–6)
	ADHD	HBQ (C,T); DISC-IV (C)	**HBQ (C, T):** Mean of *inattention* and *impulsivity* subscales; **DISC-IV:** Sum of ADHD items
	Internalizing	HBQ (C,T); DISC-IV (C)	**HBC (C):** Mean of *overanxious, depression,* and *separation anxiety* subscales; **HBC (T):** Mean of *overanxious* and *depression* subscales; **DISC-IV:** Sum of depression and anxiety disorder items

Externalizing	HBQ (C,T); DISC-IV (C)	**HBQ (C, T):** Mean of *oppositional defiant*, *conduct problems*, *overt aggression*, and *relational aggression* subscales; **DISC-IV:** Sum of oppositional defiant disorder and conduct disorder items

Notes. (C) = caregiver report. (T) = teacher report. ADHD = Attention-Deficit/Hyperactivity Disorder; ITSEA = Infant Toddler Socio and Emotional Assessment (21); DAI = Disturbances of Attachment Interview (22); PAPA = Preschool Age Psychiatric Assessment (23); HBQ = MacArthur Health and Behavior Questionnaire (24).

Figure S1. Distributions of IQ scores at each assessment wave.

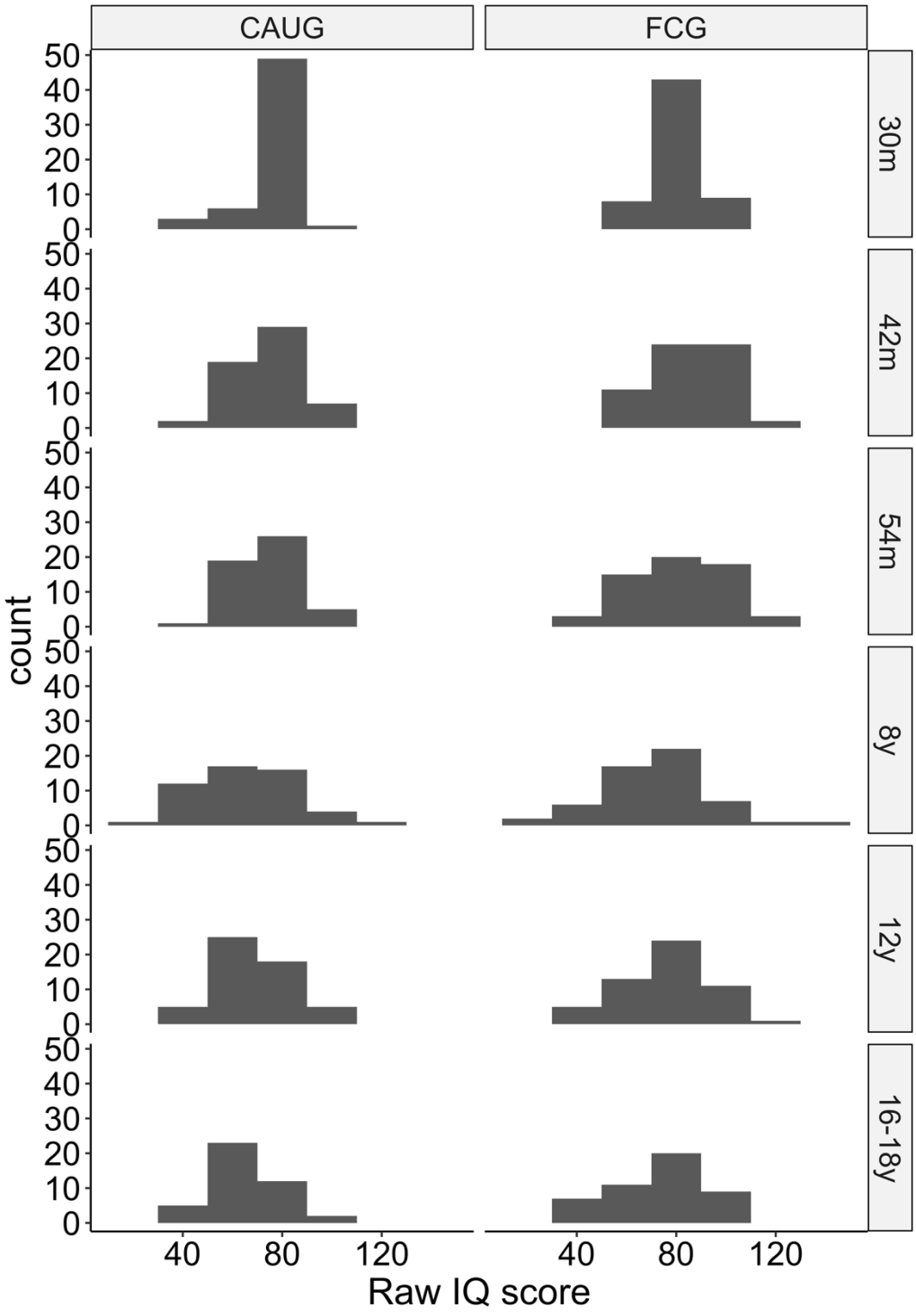


Figure S2. Distributions of EEG relative α-power at each assessment wave.

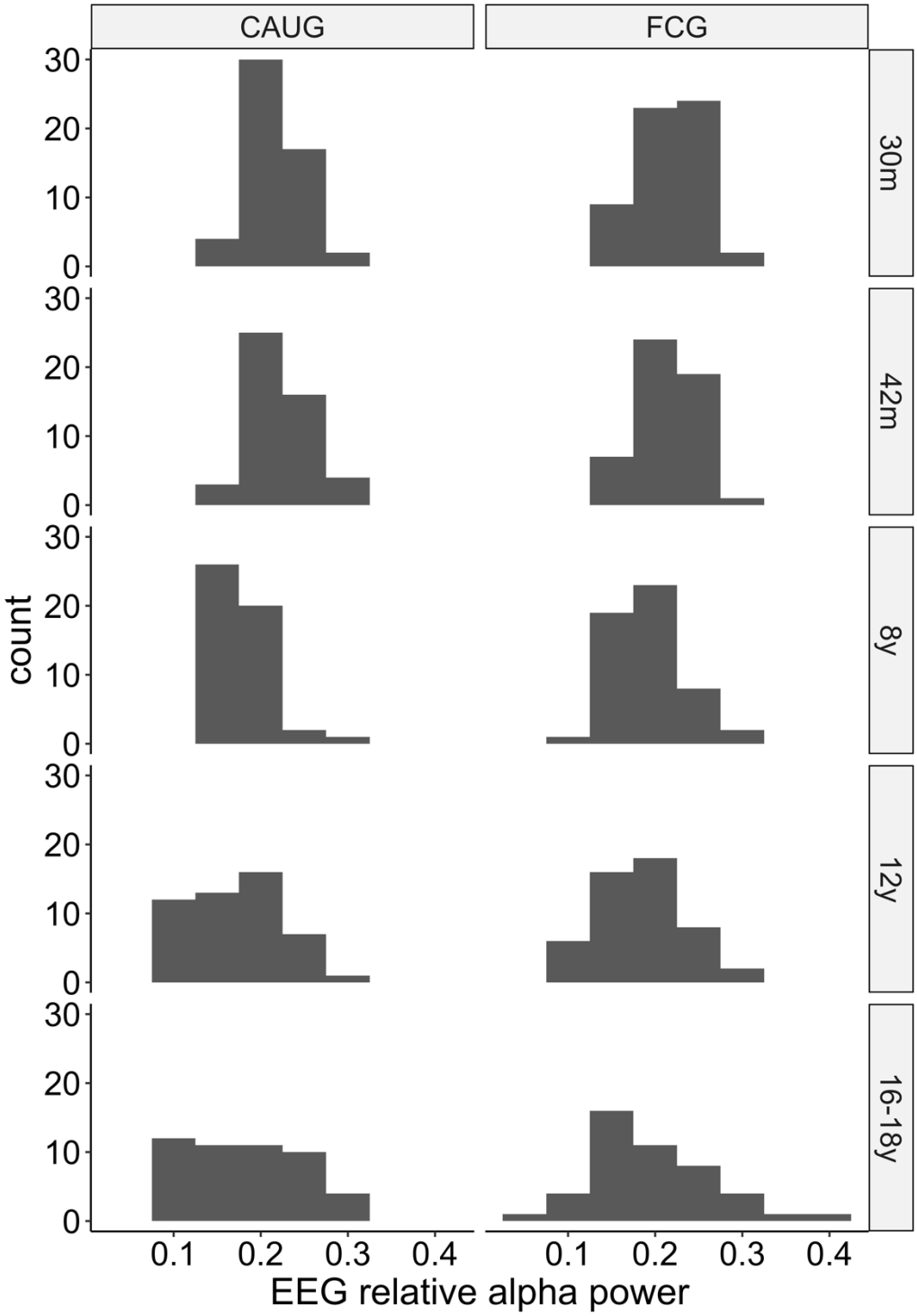


Figure S3. Distributions of height (cm) at each assessment wave.

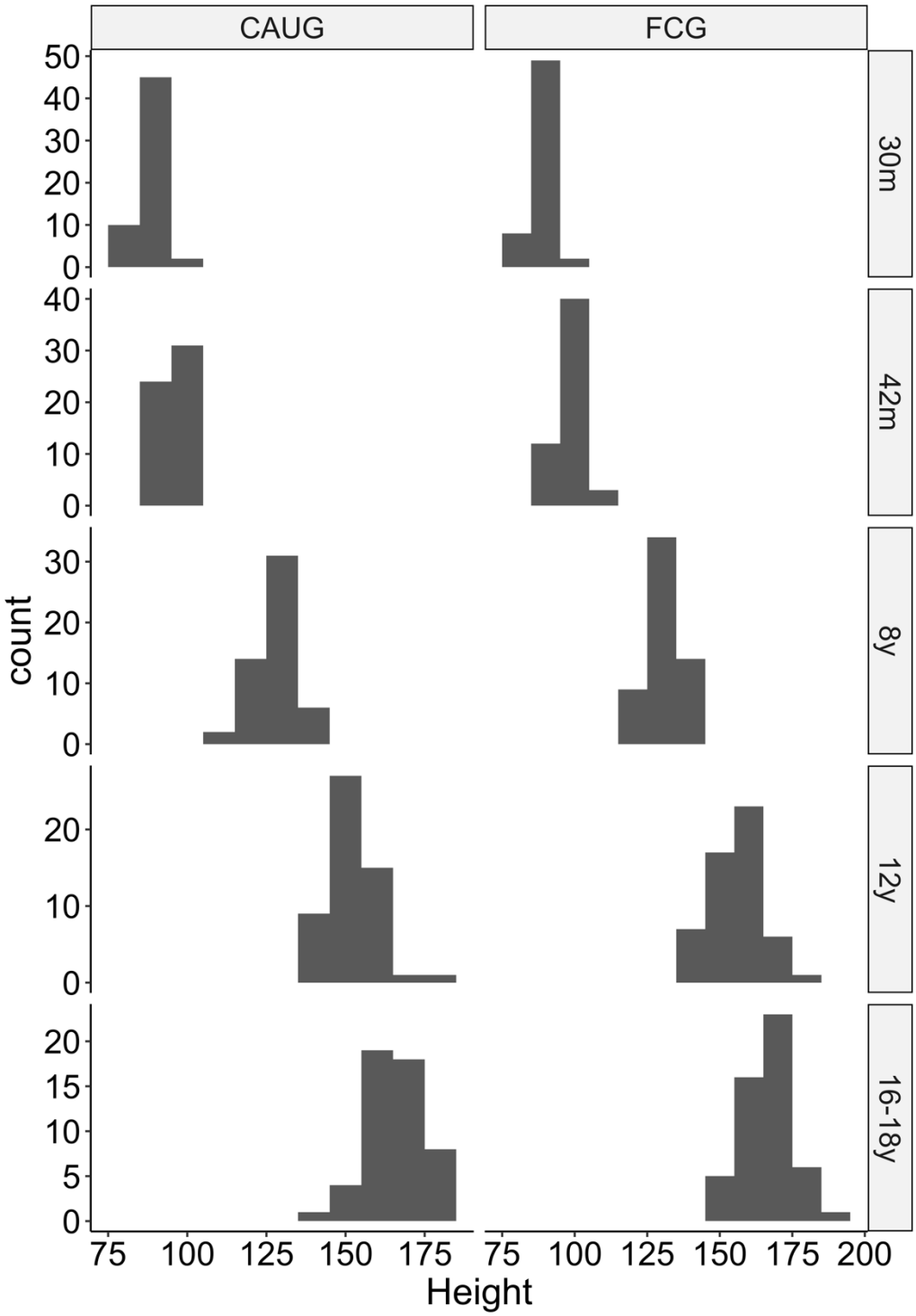


Figure S4. Distributions of weight (kg) at each assessment wave.

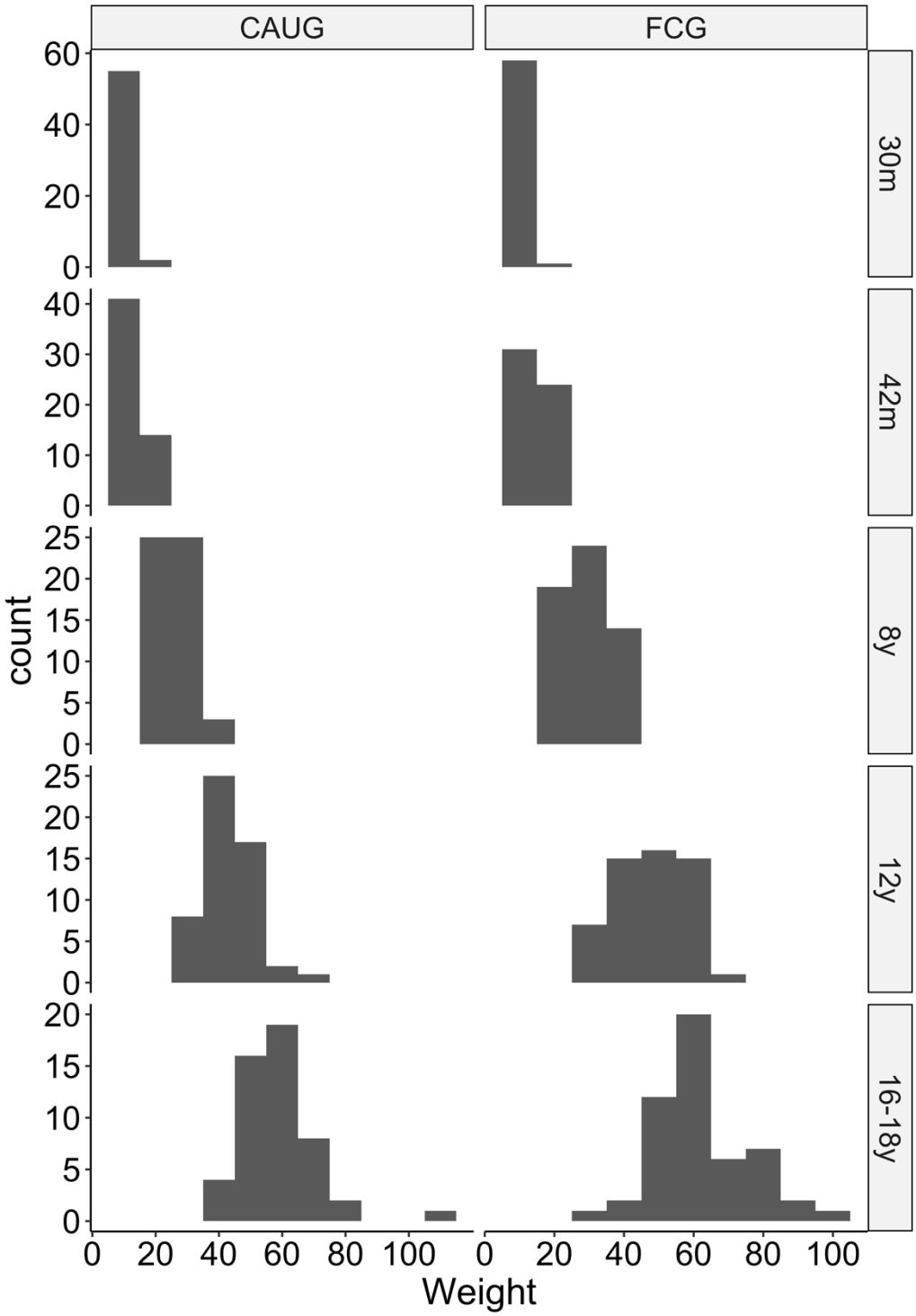


Figure S5. Distributions of head circumference (cm) at each assessment wave.

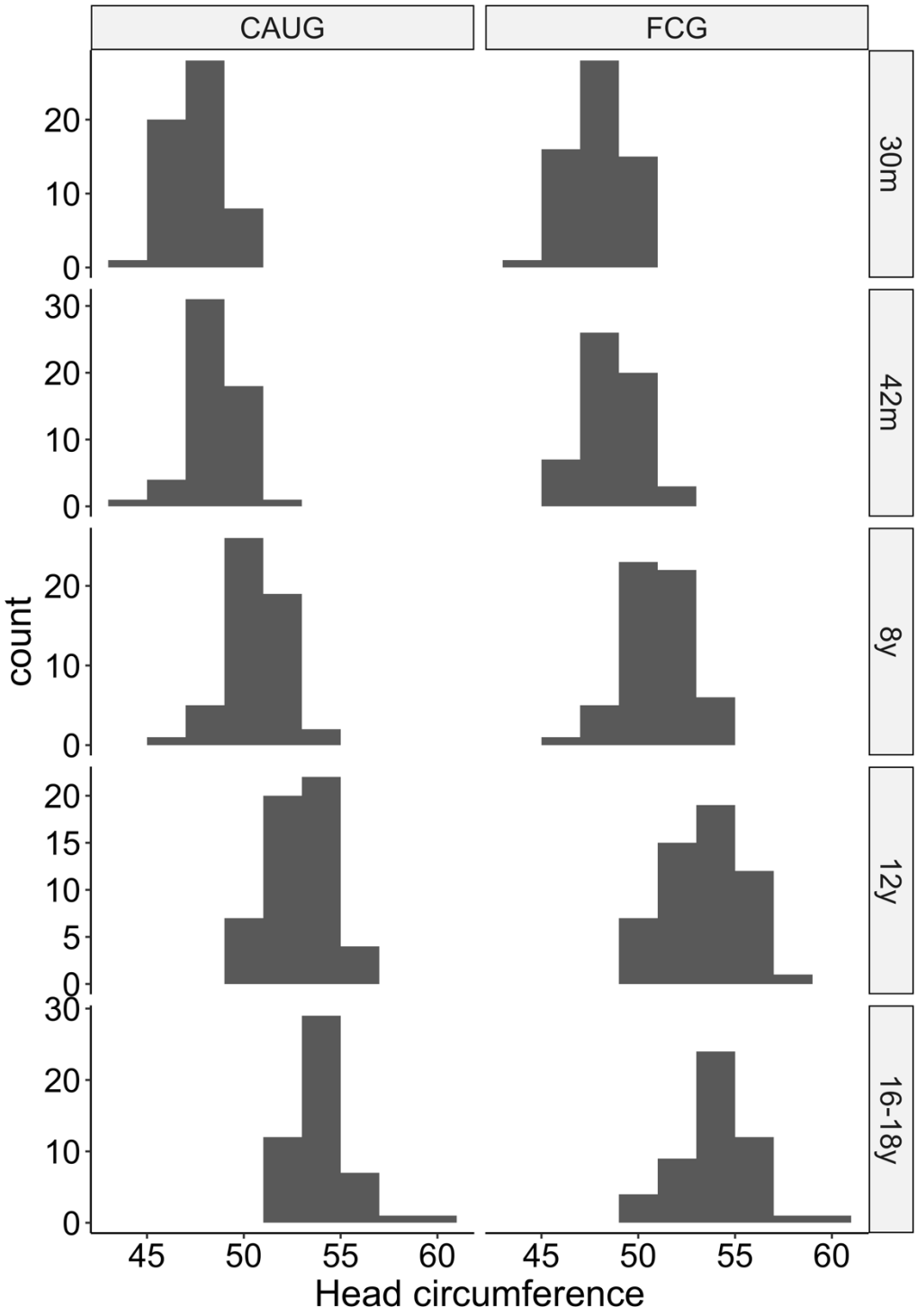


Figure S6. Distributions of disinhibited social engagement disorder symptoms at each assessment wave.

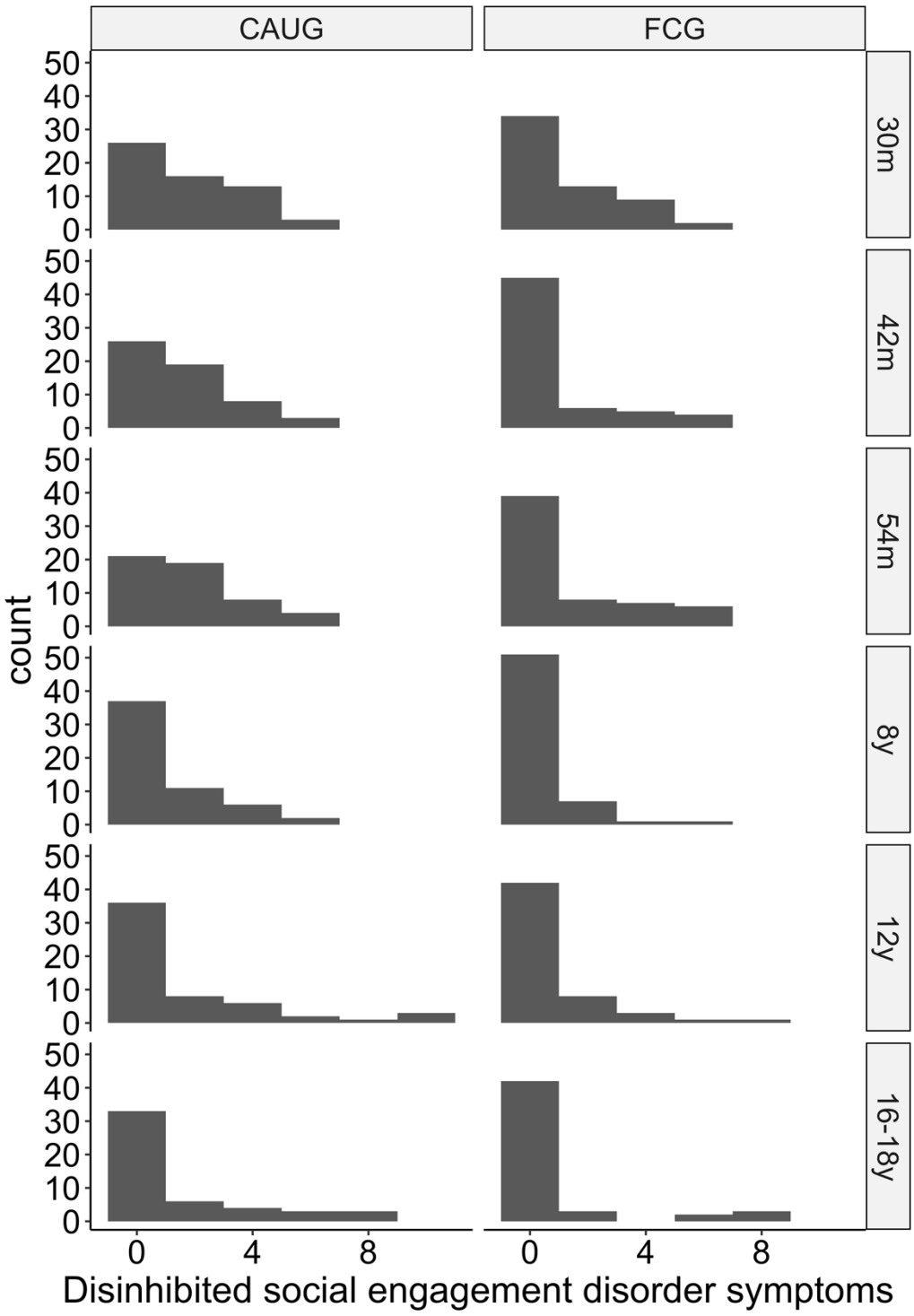


Figure S7. Distributions of reactive attachment disorder symptoms at each assessment wave.

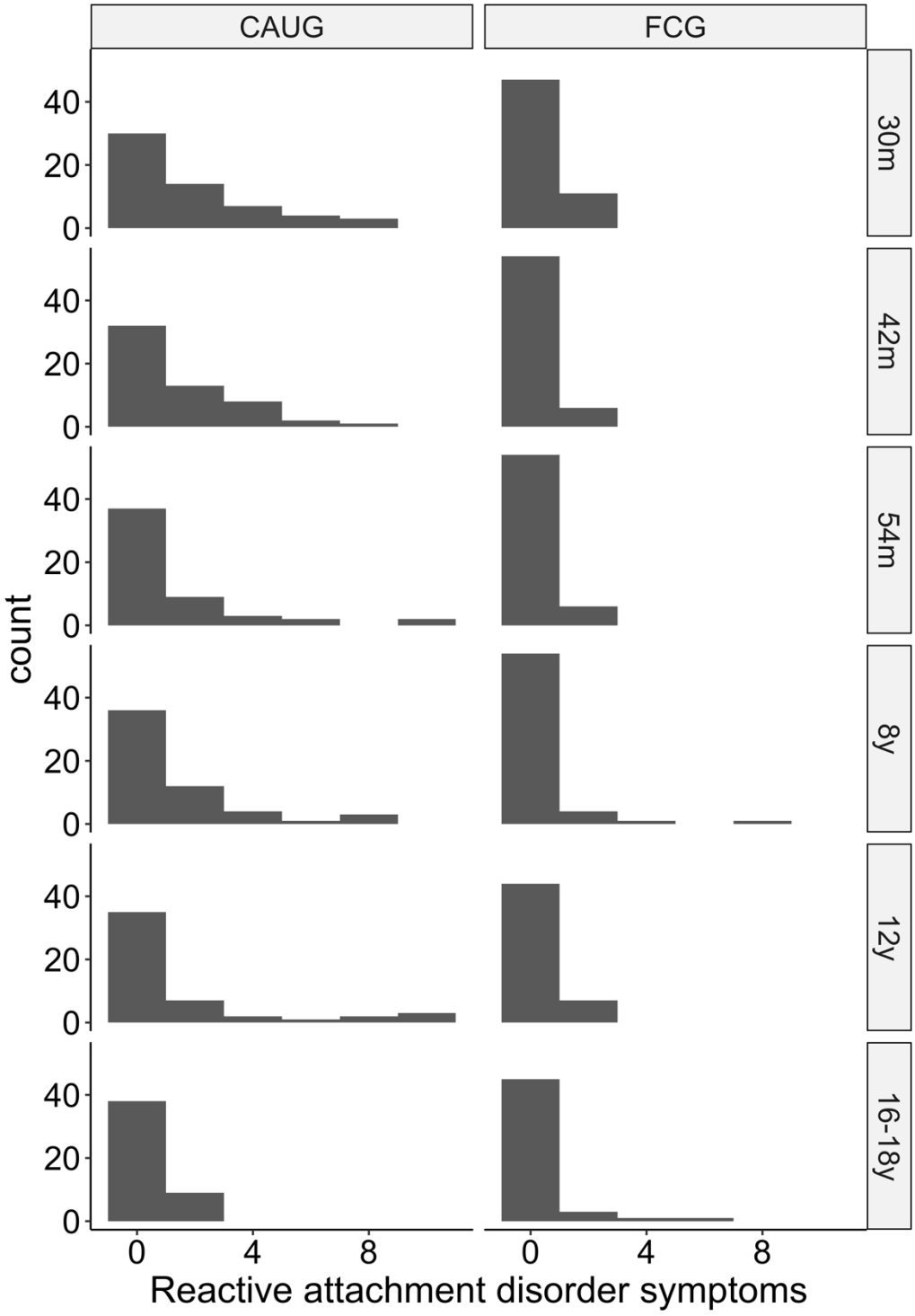
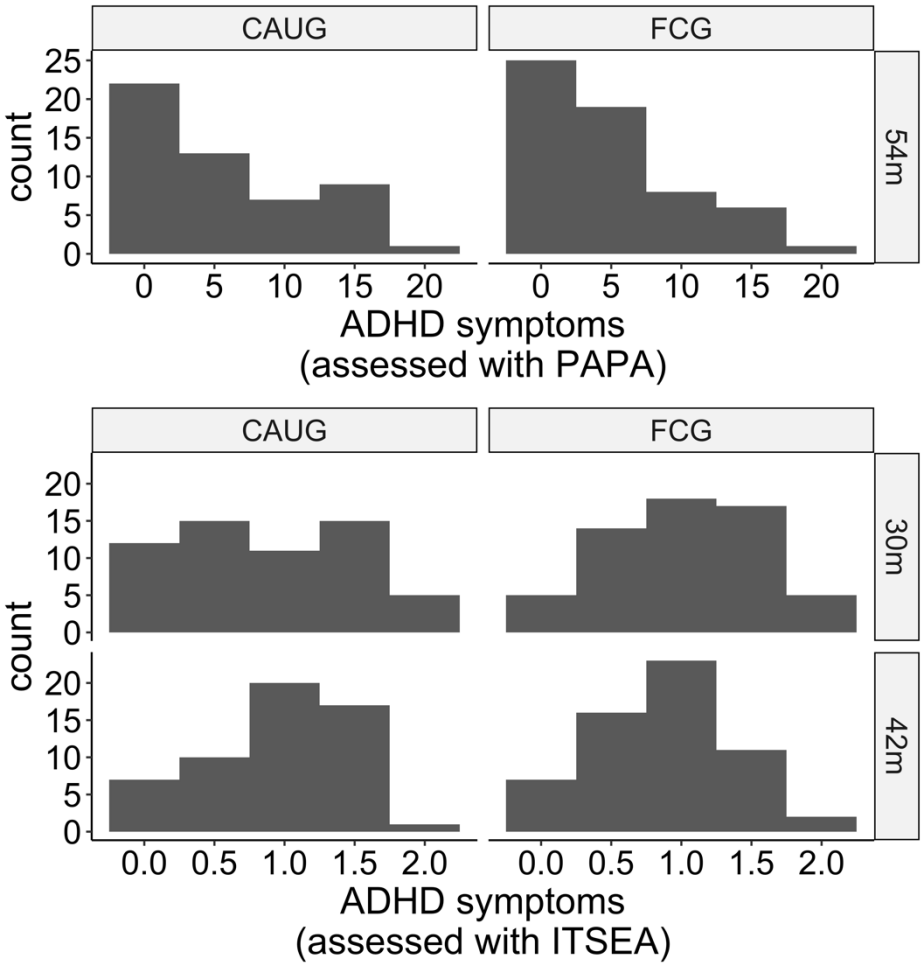
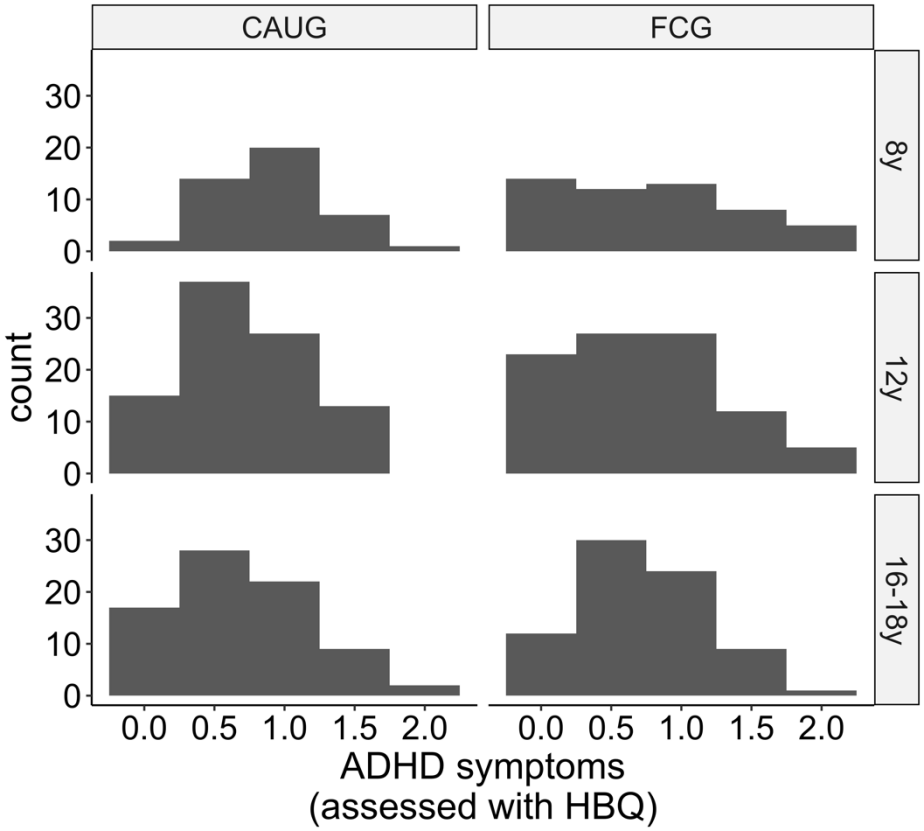


Figure S8. Distributions of ADHD symptoms at each assessment wave.




CAUG
FCG
8y
12y
16-18y
count
30
20
10
0
0.0 0.5 1.0 1.5 2.0
ADHD symptoms
(assessed with HBQ)

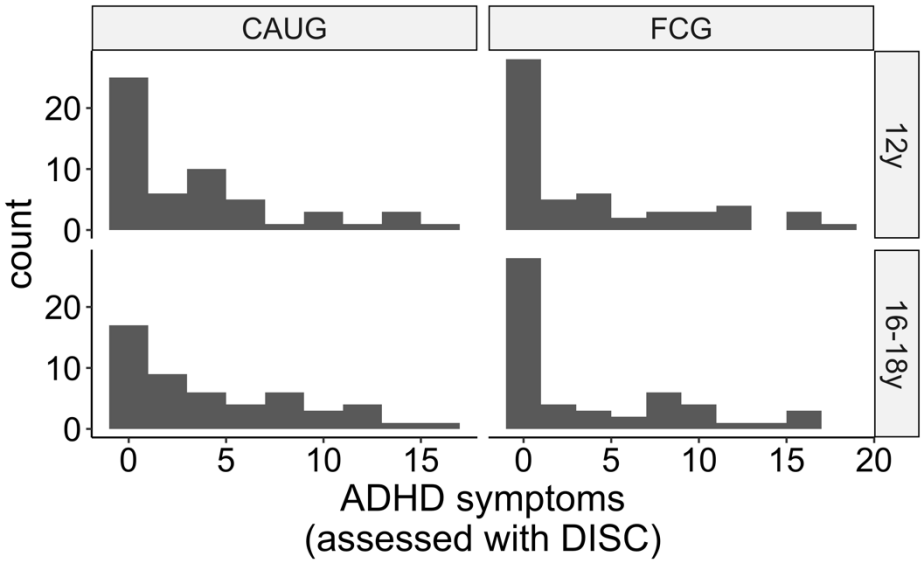
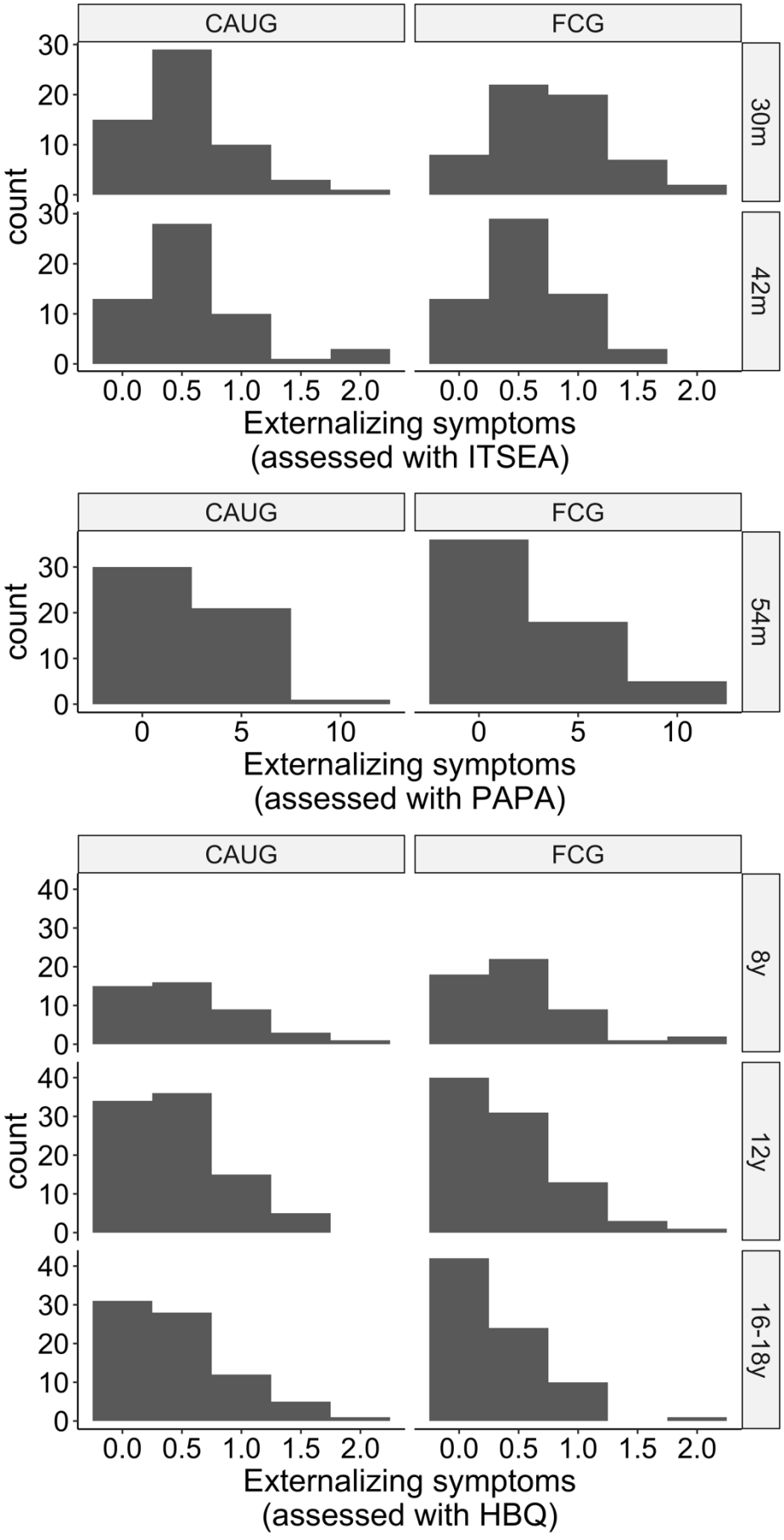

CAUG
FCG
12y
16-18y
count
20
10
0
0 5 10 15
0 5 10 15 20
ADHD symptoms
(assessed with DISC)

Figure S9. Distributions of externalizing symptoms at each assessment wave.



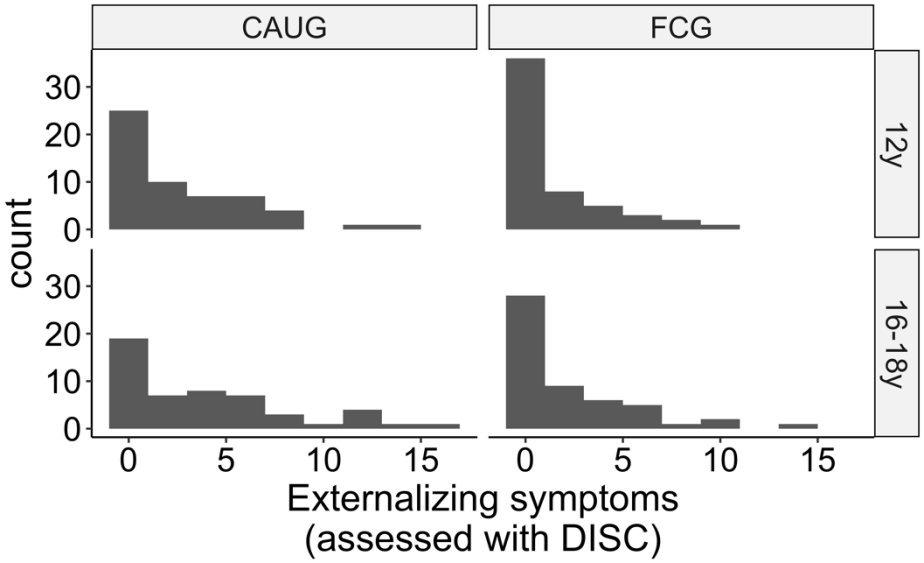
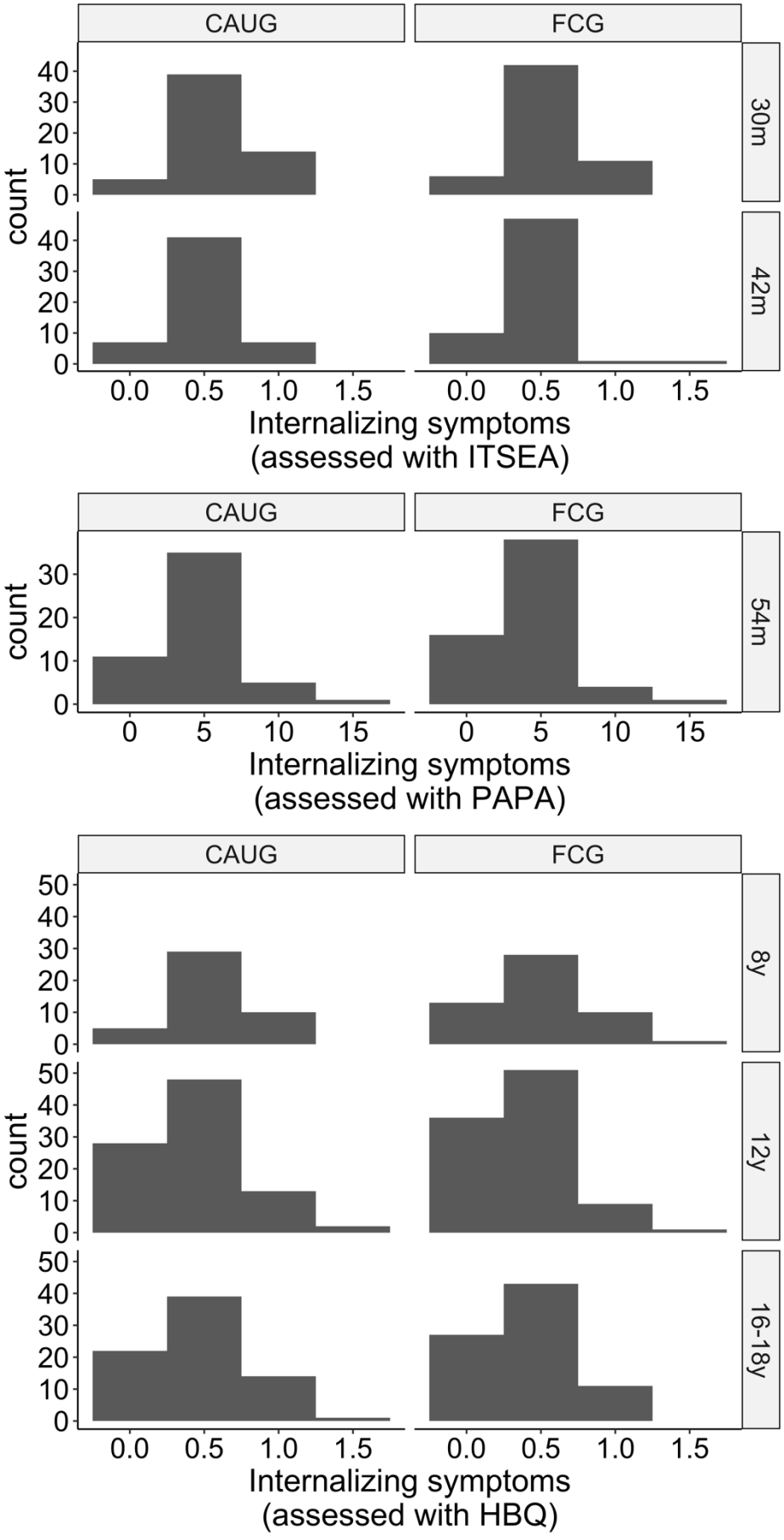
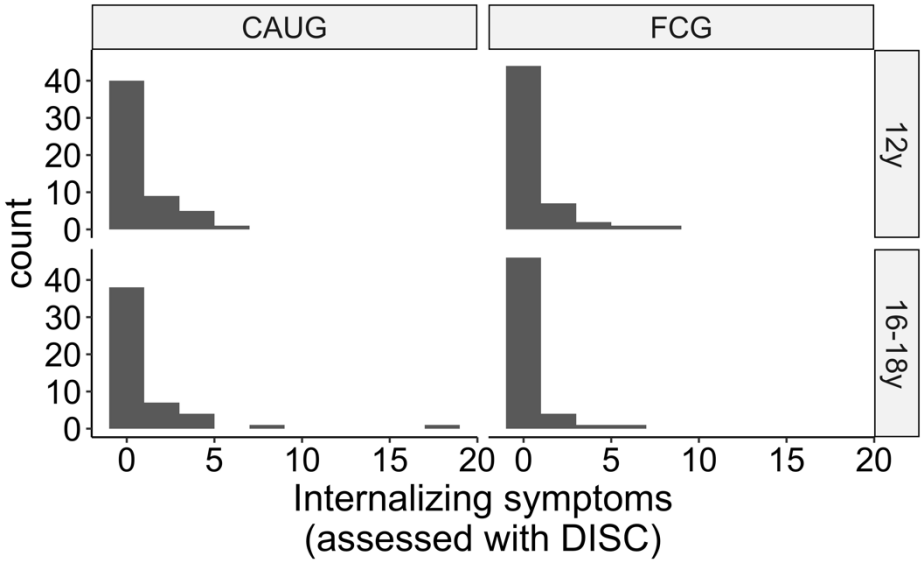

CAUG
FCG
12y
16-18y
count
30
20
10
0
0
5
10
15
Externalizing symptoms
(assessed with DISC)

Figure S10. Distributions of internalizing symptoms at each assessment wave.

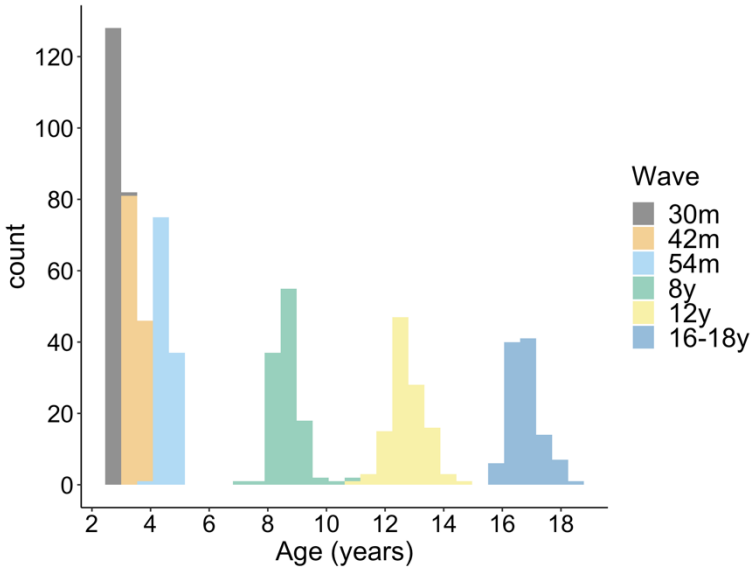



CAUG
FCG
12y
16-18y
count
40
30
20
10
0
0
5
10
15
20
Internalizing symptoms
(assessed with DISC)

Age of assessment

Given variation in age at each wave and for each measurement at each wave (see Figure S11), we modeled participant-specific age at each measurement occasion in all analyses. The age at assessment was missing for 45 observations across 22 participants. We imputed missing age values with the mean age for the participant's other assessments at that wave.

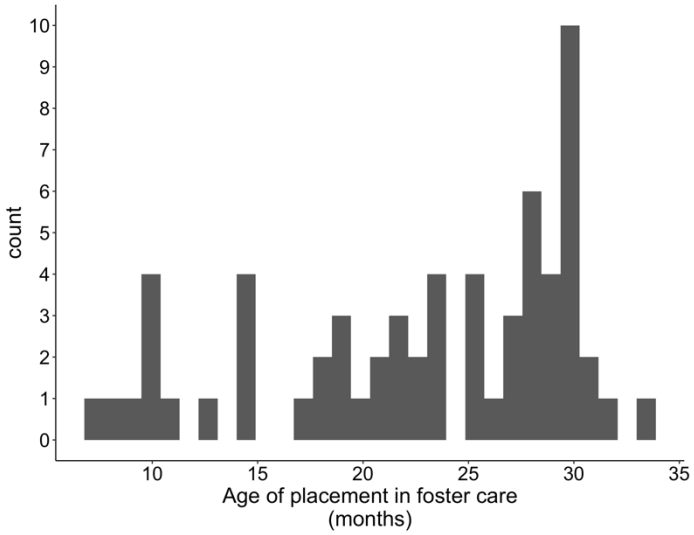
Figure S11. Distributions of mean age across assessments at each wave



Age of placement in foster care

Age of placement in foster corresponds to the age in months for each child in the FCG when they were first placed with a study-sponsored foster care family (see distribution in Figure S12).

Figure S12. Distribution of age of placement in foster care among children randomized to foster care.



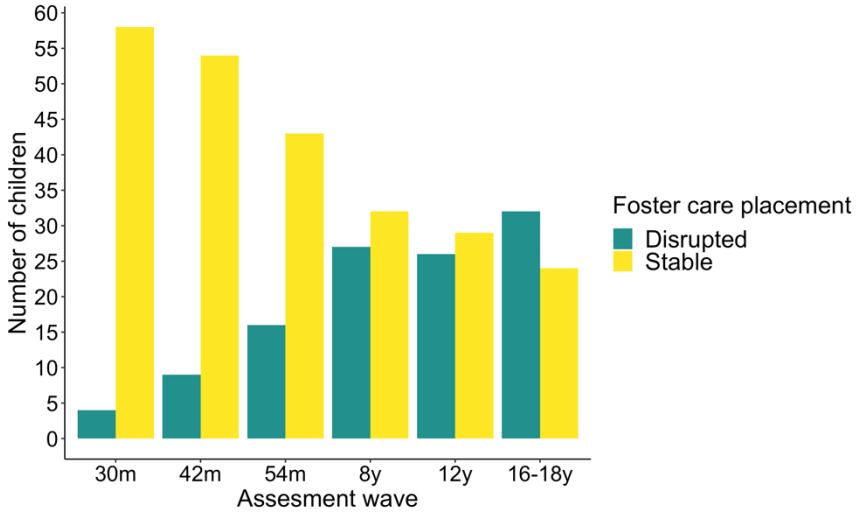
Stability of foster care placement

In previous analyses from the BEIP, the effects of placement stability on the outcomes of FCG children have been examined cross-sectionally (7, 25, 26). Specifically, for a given timepoint, all children who had experienced at least one disruption from their original study-sponsored foster care family by that timepoint were categorized as "disrupted" whereas all children who had remained with this family through that timepoint were categorized as "stable." Given that the current analyses examined children's outcomes across development, we instead treated placement stability as a time-varying covariate. For each assessment, FCG children were identified as "stable" if they remained with their original foster family at that assessment or "disrupted" if they had experienced at least one placement disruption by that assessment. Thus, as displayed

in Figure S13, across time, more children experienced placement disruptions and the number of children categorized as "disrupted" increased as the number of children categorized as "stable" decreased.

We examined whether baseline measures were associated with stability of placement within the FCG at the 8-, 12-, and 16–18-year assessment waves. We did not test associations with stability at the 30-, 42-, or 54-month waves given small cell sizes within the "disrupted" subgroup at these timepoints. Specifically, in a series of Welch's *t*-tests, we tested subgroup (stable vs. disrupted) differences in baseline symptoms of reactive attachment disorder, disinhibited social engagement disorder, ADHD, and externalizing and internalizing symptoms, and in baseline measures of IQ, head circumference, height, and weight. None of these baseline measures were associated with the stability of placement at the 8-year wave (*t*-values < |1.42|, *p*-values > .16), at the 12-year wave (*t*-values < |1.05|, *p*-values > .30), or at the 16–18-year wave (*t*-values < |0.76|, *p*-values > .45).

Figure S13. Number of children categorized as having disrupted or stable foster care placements at each assessment wave.



Missing data

Data was missing in the current analyses for several reasons, including 1) discontinuing participation (i.e., a participant withdrew from the study and subsequently never participated in additional follow-up waves), 2) missing (i.e., skipping) a follow-up wave but subsequently participating again, *or* 3) completing some assessments in a given wave but not others (e.g., completing the IQ and physical growth assessments but missing EEG).

Given these various ways data could be missing, we analyzed missingness by creating a continuous variable reflecting the proportion of all possible assessments across the six follow-up waves that were missing from each participant. Out of the 68 possible assessments that a given participant could complete across all waves, participants were missing an average of 20% (i.e., completed an average of 55 assessments).

First, we examined the association between intervention group (FCG vs. CAUG) and missingness. There was no association between missingness and intervention group (Welch's $t(126.18)=1.03$, $p=.306$).

Second, we examined whether baseline measures were differentially associated with missingness by intervention group. Specifically, in separate OLS regression models, we regressed the missingness variable onto interactions between intervention group and each of the measures collected at baseline (symptoms of reactive attachment disorder, disinhibited social engagement disorder, and ADHD; externalizing and internalizing symptoms; IQ; and measures of head circumference, height, and weight). In none of these regressions was there a significant interaction between the baseline measure and intervention group (*p*-values > .15). These results suggest that baseline characteristics were not differentially associated with missingness by intervention group and that there is minimal bias in the causal effects due to participants with poorer or better functioning within a group being more likely to discontinue or miss assessments.

Results

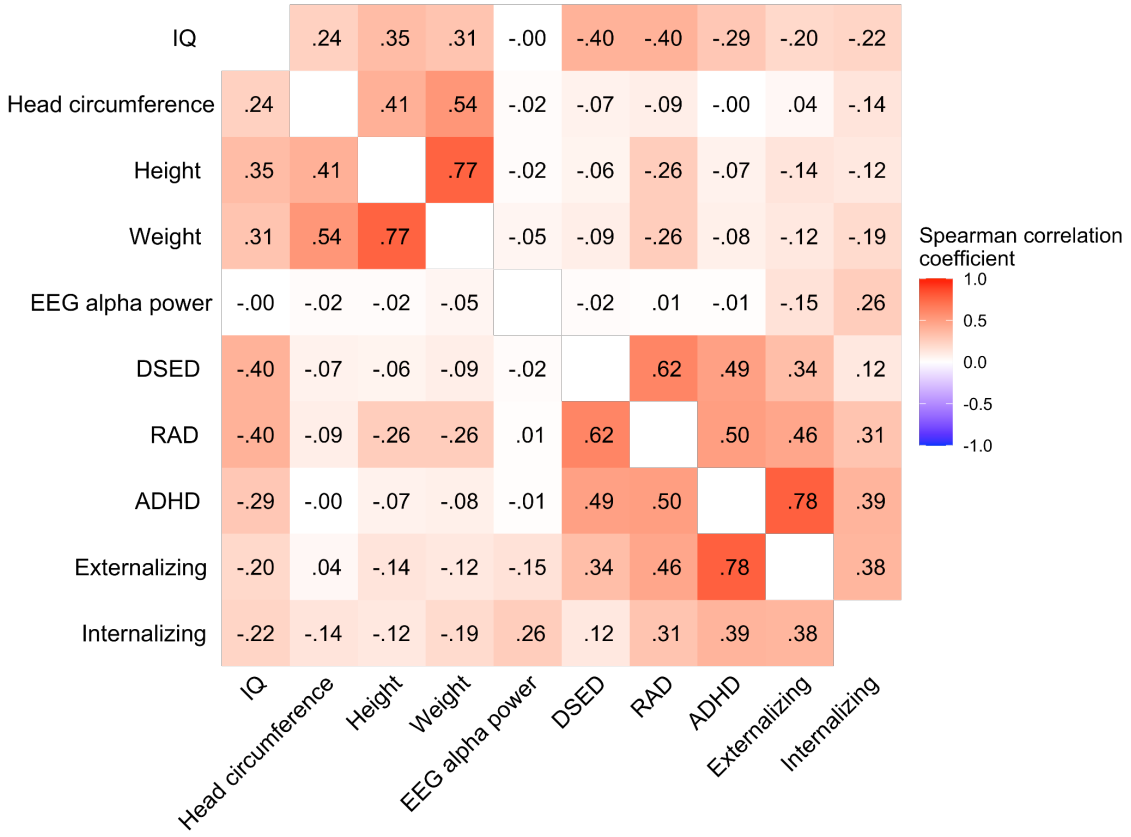
Descriptive statistics

We present descriptive statistics of the measures in each of the cognitive, physical, and neural domains and for each type of psychopathology in Table S4. We presented Spearman's correlation coefficients for the associations among standardized outcome scores averaged across waves for each outcome domain and type of psychopathology in Figure S14.

Table S4. Descriptive statistics for outcomes averaged across assessment waves. For all forms of psychopathology, scores represent symptom severity. IQ, physical growth, EEG α-power, disinhibited social engagement disorder, and reactive attachment disorder were measured in the same units across ages of assessment. Measures of ADHD, externalizing, and internalizing symptoms differed across development (ITSEA = Infant Toddler Social Emotional Assessment; PAPA = Preschool Psychiatric Assessment; HBQ = MacArthur Health and Behavior Questionnaire, DISC = Diagnostic Interview Schedule for Children). Means are raw values averaged across assessment waves. SD = standard deviation.

Outcome	Foster care		Care as usual	
	Mean	**SD**	**Mean**	**SD**
Cognitive, physical, and neural				
IQ	78.89	20.67	71.97	15.21
Physical growth				
Head circumference (cm)	51.02	3.01	50.74	2.90
Height (cm)	126.97	31.24	125.01	31.13
Weight (kg)	32.59	20.67	30.01	18.68
EEG relative alpha power	0.20	0.05	0.19	0.05
Psychopathology				
Disinhibited social engagement disorder	1.15	1.87	1.83	2.12
Reactive attachment disorder	0.48	0.97	1.59	2.29
ADHD				
ITSEA	0.96	0.52	0.92	0.56
PAPA	4.90	4.82	5.79	5.49
HBQ	0.75	0.52	0.75	0.45
DISC	4.13	5.32	4.27	4.61
Externalizing				
ITSEA	0.68	0.43	0.56	0.43
PAPA	2.36	2.86	2.44	2.48
HBQ	0.41	0.39	0.49	0.43
DISC	2.16	2.91	3.74	3.92
Internalizing				
ITSEA	0.51	0.22	0.56	0.23
PAPA	4.10	2.69	4.56	2.53
HBQ	0.43	0.29	0.47	0.29
DISC	0.64	1.51	1.17	2.34

Figure S14. Correlations among standardized outcome scores averaged across assessment waves.



Sources of variation in the effects of the intervention on children's outcomes

Cognitive, physical, and neural

Results of simple effect analyses for the two-way interaction between group and outcome domain indicated that the effects of the intervention on IQ and physical growth were significantly larger than on EEG relative alpha power (IQ vs. frontal EEG alpha power: β=0.33, 95% CI [0.14, 0.52], p<.001; physical growth vs. EEG relative alpha power: β=0.21, 95% CI [0.06, 0.38], p=.006); there was no significant difference between the effects on IQ and physical growth.

There was a significant three-way interaction between intervention group, outcome domain, and sex (F(2, 2656.58)=12.11, p<.001). We present the simple effects of intervention group for each outcome domain and sex assigned at birth in Table S5. Girls randomized to foster care had significantly higher IQ scores than their peers randomized to care as usual whereas boys did not; the effect of the intervention on IQ was significantly larger in girls than in boys (β=0.49, 95% CI [0.01, 0.94], p=.050). Both boys and girls randomized to foster care had greater physical growth than their peers randomized to care as usual and this effect was similar in magnitude in boys and girls. Boys randomized to foster care had greater EEG alpha power than their peers randomized to care as usual whereas girls randomized to foster care had lower frontal EEG alpha power. Although the effect of the intervention on EEG alpha power was significantly larger in boys than in girls (β=0.48, 95% CI [0.01, 0.92], p=.044), it was not significant within either boys or girls.

Table S5. Differences in cognitive, physical, and neural outcomes between children randomized to the foster care intervention compared to children randomized to care as usual based on outcome domain and sex assigned at birth. β = standardized mean group differences, adjusting for covariates. β-values, 95% CIs, and *p*-values are bootstrapped (1,000 iterations).

Domain	Sex	β	95% CI	*p*
IQ	Male	0.16	-0.15, 0.48	.346
	Female	0.65	0.33, 0.99	<.001
Physical size	Male	0.29	<0.01, 0.57	.052
	Female	0.27	-0.03, 0.55	.070
EEG alpha power	Male	0.29	-0.04, 0.62	.084
	Female	-0.19	-0.51, 0.14	.250

Psychopathology

Results of simple effect analyses for the two-way interaction between group and type of psychopathology indicated that the effects of the intervention on disinhibited social engagement disorder and reactive attachment disorder were significantly larger than the effects on ADHD (disinhibited social engagement disorder vs. ADHD: β=0.31, 95% CI [0.13, 0.49], p<.001; reactive attachment disorder vs. ADHD: β=0.56, 95% CI [0.39, 0.72], p<.001) and externalizing symptoms (disinhibited social engagement disorder vs. externalizing: β=0.21, 95% CI [0.04, 0.38], p=.016; reactive attachment disorder vs. externalizing: β=0.46, 95% CI [0.28, 0.64], p<.001). Further, the effect of the intervention on reactive attachment disorder symptoms was significantly larger than the effects on internalizing (β=0.39, 95% CI [0.22, 0.56], p<.001) and disinhibited social engagement disorder symptoms (β=0.26, 95% CI [0.05, 0.44], p=.008). Finally, the effect of the intervention on internalizing symptoms was significantly larger than the effect on ADHD symptoms (β=0.17, 95% CI [0.01, 0.33], p=.032). There were no significant differences between the effects of the intervention on disinhibited social engagement disorder and internalizing symptoms, ADHD and externalizing symptoms, or between internalizing and externalizing symptoms.

There was a significant three-way interaction between intervention group, type of psychopathology, and age of assessment (F[4, 4149.00], 5.91, p<.001). We present the simple effects of intervention group for each type of psychopathology centered at ages 3.5, 8, and 16 years in Table S6. Whereas effect sizes for symptoms of disinhibited social engagement disorder, ADHD, and internalizing symptoms were similar across ages of assessment, the effects of the intervention varied significantly based on age of assessment for reactive attachment disorder (β=0.03, 95% CI [0.01, 0.06], p=.020) and externalizing symptoms (β=-0.05, 95% CI [-0.07, -0.03], p<.001). Specifically, the effect on reactive attachment disorder symptoms was significant at every age, but this effect was smaller in adolescence compared to at younger ages; in contrast, the effect on externalizing symptoms was small and not statistically significant at younger ages but medium in magnitude and statistically significant in adolescence.

Table S6. Differences in symptoms of psychopathology between children randomized to the foster care intervention compared to children randomized to care as usual based on type of psychopathology and age of assessment. β = standardized mean group differences, adjusting for covariates. β-values, 95% CIs, and *p*-values are bootstrapped (1,000 iterations).

Type of psychopathology	Assessment age (years)	β	95% CI	*p*
Disinhibited social engagement disorder	3.5	-0.36	-0.58, -0.13	.002
	8	-0.36	-0.56, -0.15	<.001
	16	-0.35	-0.65, -0.04	.022
Reactive attachment disorder	3.5	-0.75	-0.99, -0.52	<.001
	8	-0.60	-0.80, -0.38	<.001
	16	-0.32	-0.64, -0.02	.036
ADHD	3.5	-0.03	-0.26, 0.21	.814
	8	-0.04	-0.23, 0.15	.688
	16	-0.08	-0.31, 0.15	.546
Externalizing	3.5	0.16	-0.07, 0.39	.162
	8	-0.05	-0.24, 0.15	.592
	16	-0.42	-0.64, -0.21	<.001
Internalizing	3.5	-0.20	-0.44, 0.05	.112
	8	-0.21	-0.41, -0.01	.042
	16	-0.24	-0.48, <0.01	.052

Sources of variation among children randomized to foster care: timing and stability of placement

Cognitive, physical, and neural

Results of simple effect analyses for the two-way interaction between age of placement and outcome domain indicated that the effect of age of placement on IQ was significantly larger than on physical size (β=0.18, 95% CI [0.07, 0.29], p<.001), but this effect was not significantly different between IQ and EEG alpha power or between EEG alpha power and physical size.

We probed the two-way interaction between age of placement and age of assessment by examining the simple effects (displayed in Table S7) of age of placement on cognitive, physical, and neural outcomes when age of assessment was centered at 3.5, 8, and 16 years. Age of placement in foster care was significantly negatively associated with children's cognitive, physical, and neural outcomes at younger ages, but not in adolescence.

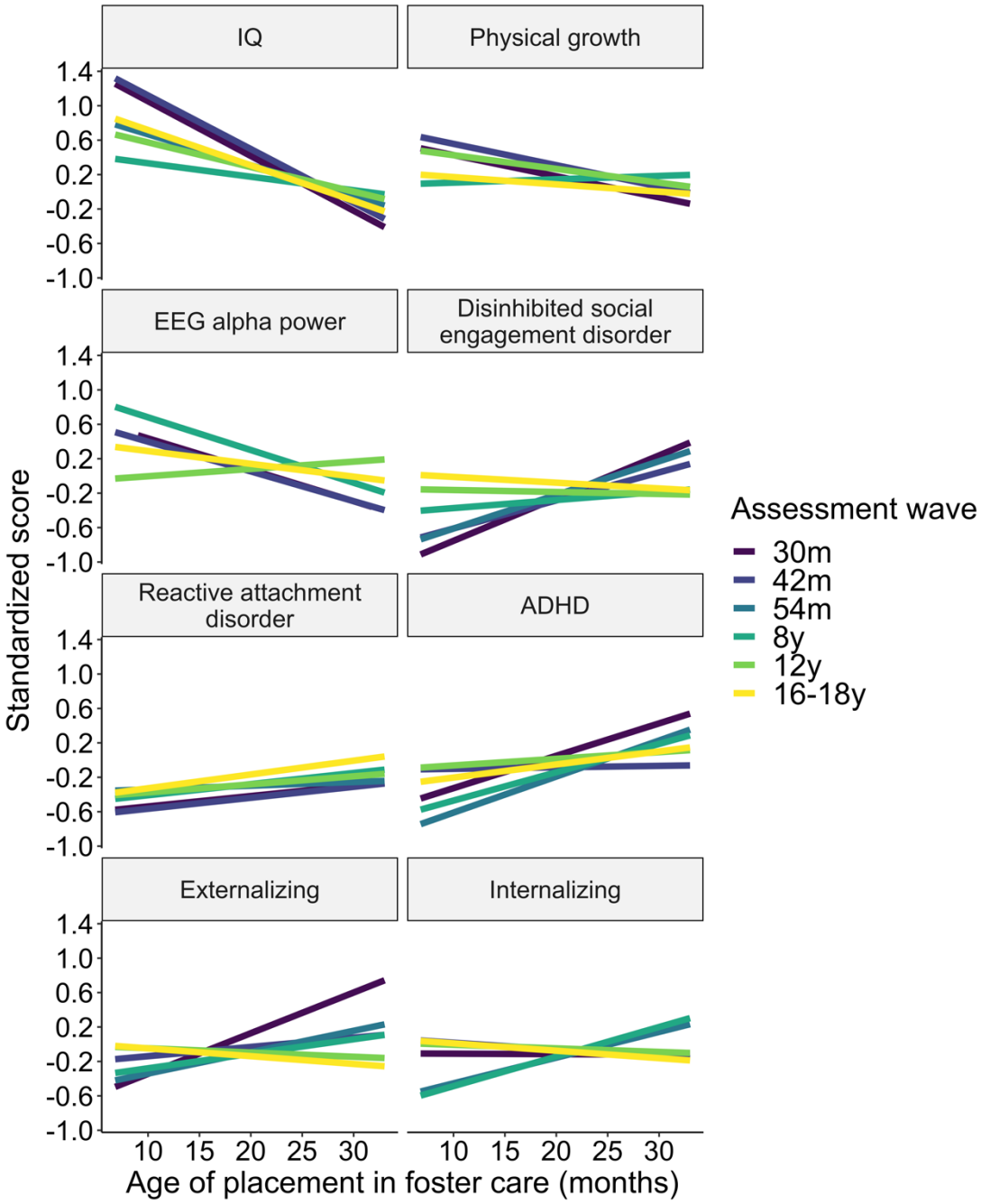
Psychopathology

We probed the two-way interaction between age of placement and age of assessment by examining the simple effects (displayed in Table S7) of age of placement on symptoms of psychopathology when age of assessment was centered at 3.5, 8, and 16 years. Age of placement in foster care was positively associated with children's symptoms in early childhood, but was not associated with children's symptoms in middle childhood or adolescence.

Table S7. Associations of age of placement with cognitive, physical, and neural outcomes and symptoms of psychopathology based on age of assessment.

	Assessment age (years)	β	95% CI	*p*
Cognitive, physical, and neural	3.5	-0.23	-0.40, -0.08	.002
	8	-0.18	-0.35, -0.03	.018
	16	-0.09	-0.27, 0.08	.306
Psychopathology	3.5	0.14	0.01, .027	.032
	8	0.08	-0.05, 0.20	.236
	16	-0.04	-0.17, 0.09	.530

Figure S15. Among children randomized to the foster care intervention, associations of age of placement with outcomes for each domain by assessment wave. Age was analyzed continuously but is grouped by assessment wave for purposes of visualization. Each regression line reflects the linear association between age of placement into foster care and standardized outcome scores at that assessment wave.



Supplementary analyses using an extended random effects structure

We ran supplementary analyses using an extended random effects structure for our primary models presented in the main manuscript. Specifically, to account for potential non-independence of outcome scores within a given domain for a given individual, we included in these supplementary analyses a random effect for the interaction between participant and domain.

Our conclusions remained the same using this more complex modeling strategy. There were significant overall effects of the foster care intervention on cognitive, physical, and neural outcomes (β=0.27, 95% CI[0.08, 0.46], *p*=.002), and on symptoms of psychopathology (β=-0.25, 95% CI[-0.42, -0.09]). Age of placement in foster care was significantly negatively associated with cognitive, physical, and neural outcomes (β=-0.18, 95% CI [-0.32, -0.05], *p*=.004), but was not associated with symptoms of psychopathology (β=0.06, 95% CI [-0.05, 0.18], *p*=.320). Finally, in interaction with age of assessment, stability of placement in foster care was significantly associated with cognitive, physical, and neural outcomes (β=0.17, 95% CI [0.08, 0.27], *p*<.001), and with symptoms of psychopathology (β=-0.10, 95% CI [0.18, 0.02], *p*=.018).

Supplementary analyses treating age of assessment categorically

In our primary analyses, age of assessment was treated as a continuous variable reflecting exact age at each observation. We ran supplemental analyses testing whether assessment wave—a six-level categorical variable—moderated the effect of the foster care intervention on cognitive, physical, and neural outcomes and symptoms of psychopathology. Assessment wave did not significantly interact with intervention group to predict children's cognitive, physical, and neural outcomes (*F*[5, 2679.83]=1.42, *p*=.213) or symptoms of psychopathology (*F*[5, 4192.20]=1.48, *p*=.193). Thus, consistent with findings related to exact age of assessment, the overall effects of the foster care intervention did not vary by assessment wave. Standardized scores at each assessment wave for each group, outcome domain, and type of psychopathology are depicted in Figure S16.

We also ran supplemental analyses testing whether assessment wave moderated the effect of age of placement in foster care on cognitive, physical, and neural outcomes and symptoms of psychopathology among children in the foster care group. Assessment wave significantly interacted with age of placement in foster care to explain variation in children's cognitive, physical, and neural outcomes (*F*[5, 1327.75]=2.24, *p*=.049). Simple effects analyses (presented in Table S8) indicated that age of placement in foster care was significantly negatively associated with cognitive, physical, and neural outcomes only at the 30- and 42-month assessment waves. Assessment wave also significantly interacted with age of placement in foster care to explain variation in symptoms of psychopathology (*F*[5, 2085.02]=4.26, *p*<.001). Simple effects analyses (presented in Table S8) indicate that age of placement in foster care was significantly positively associated with children's symptoms of psychopathology only at the 30- and 54-month assessment waves. These results are generally consistent with those of models presented in the main manuscript treating age as a continuous variable, in which age of placement in foster care was significantly associated with children's outcomes in early childhood but not in middle childhood or adolescence.

Finally, we ran supplemental analyses testing whether the time-varying placement stability variable was associated with children's outcomes in interaction with assessment wave. Standardized sores for children with disrupted and stable foster care placements at each assessment wave for each outcome wave and type of psychopathology are plotted in Figure S17. Placement stability was significantly associated with children's cognitive, physical, and neural outcomes in interaction with assessment wave (*F*[5, 1337.67]=2.76, *p*=.017). Simple effects analyses (presented in Table S9) indicated that placement stability was significantly associated with cognitive, physical, and neural outcomes only at the 8- and 16–18-year assessment waves. These results are generally consistent with those of the models presented in the main manuscript treating age as a continuous variable, in which placement stability was associated with children's cognitive, physical, and neural outcomes in adolescence but not in early childhood. In contrast to the model presented in the main manuscript, there was no significant overall interaction between placement stability and assessment wave to explain variation in symptoms of psychopathology (*F*[5, 2028.77]=1.24, .288). Consistent with the results presented in the main manuscript, however, simple effects analyses (presented in Table S9) indicated that placement

stability was significantly associated with symptoms of psychopathology only at the 12- and 16–18-year assessment waves.

Figure S16. Standardized scores for children randomized to the foster care intervention and care-as-usual children for each outcome domain, type of psychopathology, and assessment wave. Points are means and error bars are standard errors. Physical growth and EEG alpha power were not measured at the 54-month wave.

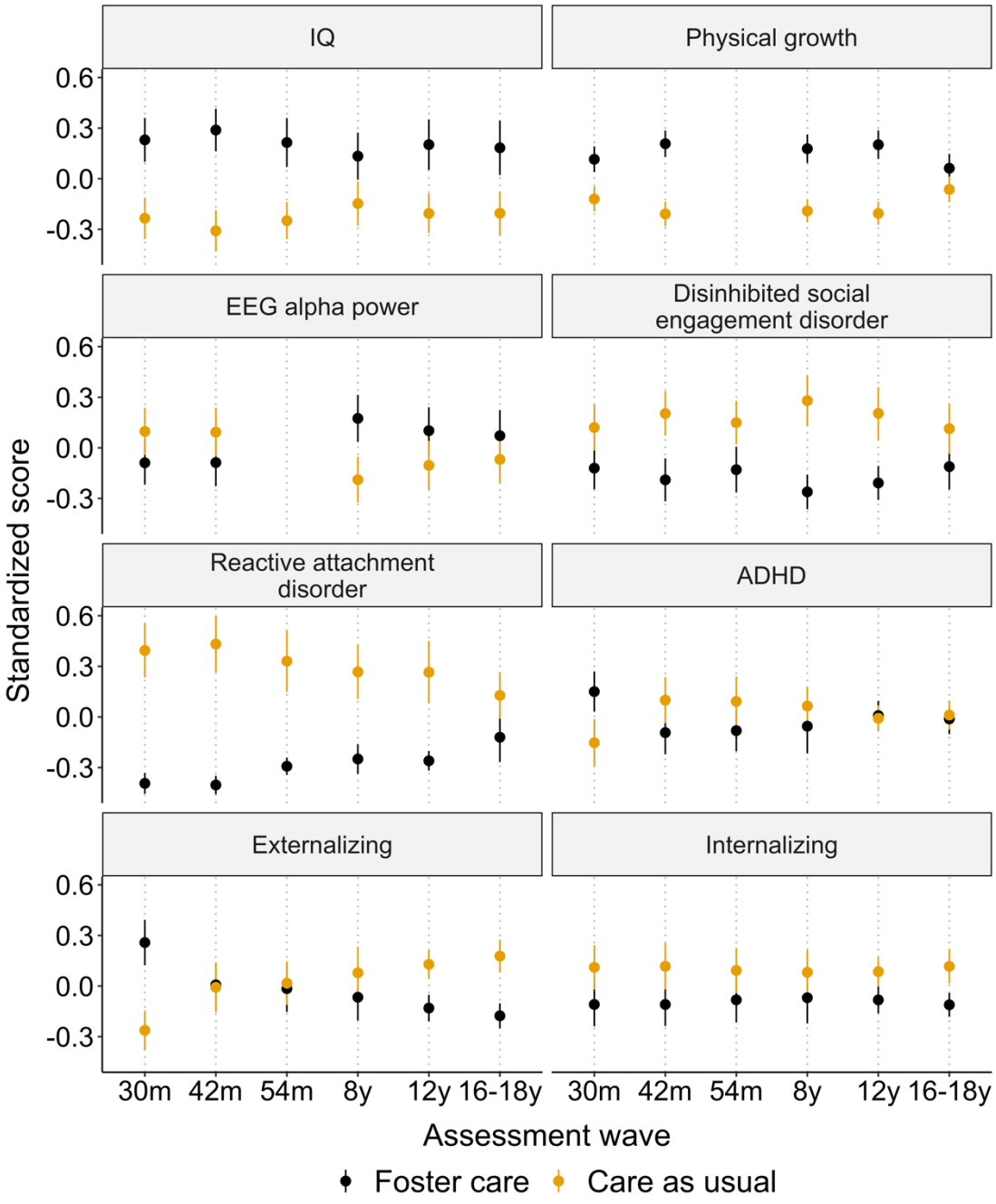


Table S8. Among children randomized to foster care, association of age of placement with cognitive, physical, and neural outcomes and symptoms of psychopathology at each assessment wave. β = association of age of placement, adjusting for covariates. β-values, 95% CIs, and *p*-values are bootstrapped (1,000 iterations).

	Assessment wave	β	95% CI	*p*
Cognitive, physical, and neural	30 months	-0.28	-0.45, -0.08	.002
	42 months	-0.28	-0.46, -0.10	.006
	54 months	-0.25	-0.51, 0.02	.080
	8 years	-0.08	-0.26, 0.10	.340
	12 years	-0.11	-0.29, 0.06	.220
	16 years	-0.13	-0.33, 0.06	.142
Psychopathology	30 months	0.17	0.02, 0.32	.020
	42 months	0.05	-0.11, 0.19	.568
	54 months	0.16	0.01, 0.30	.034
	8 years	0.12	-0.03, 0.27	.120
	12 years	<0.01	-0.15, 0.12	.954
	16 years	-0.05	-0.20, 0.08	.404

Figure S17. Among children randomized to foster care, standardized scores for children with disrupted and stable placements at each assessment wave for each outcome domain and type of psychopathology. Points are means and error bars are standard errors. Physical growth and EEG alpha power were not measured at the 54-month wave.

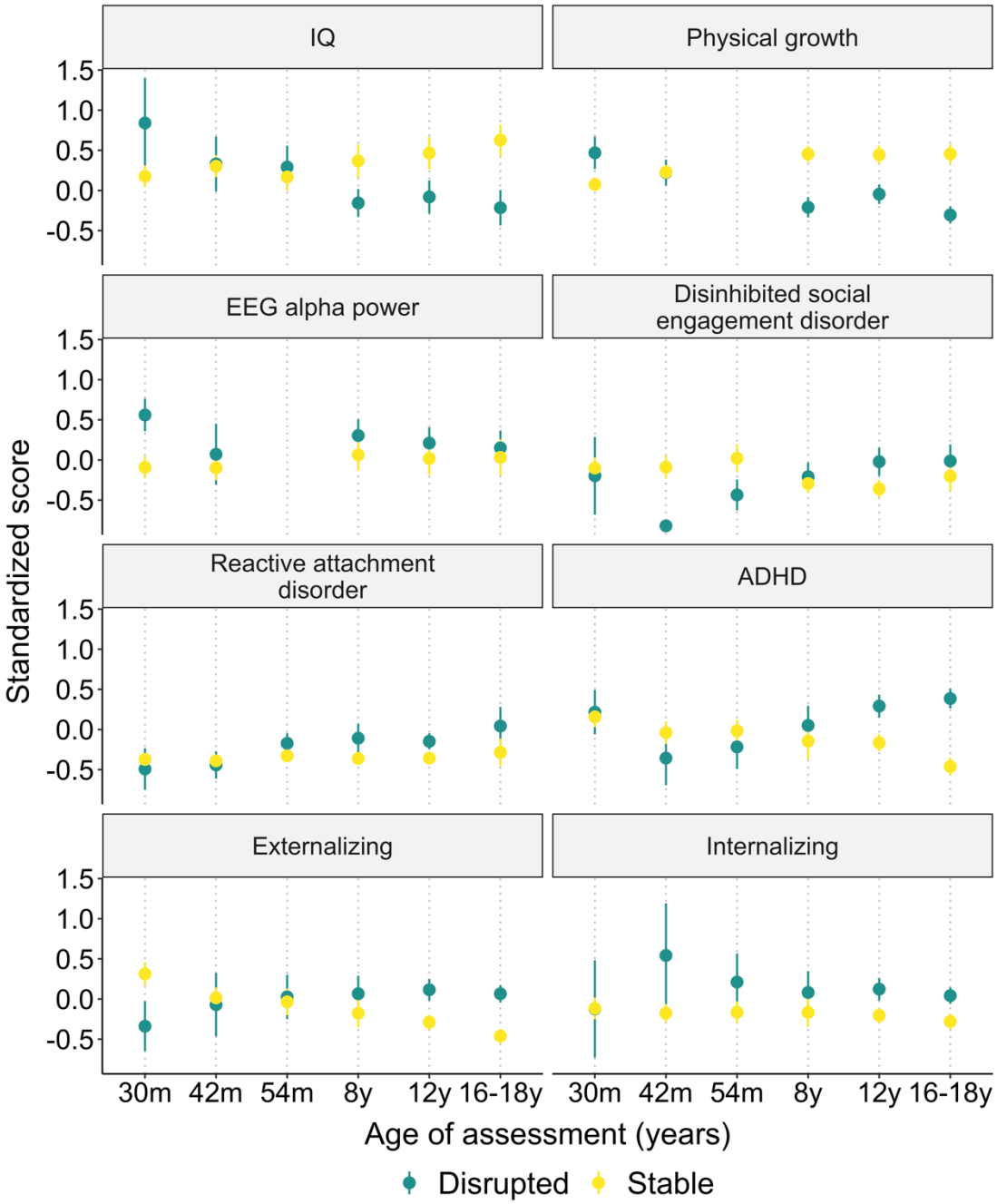


Table S9. Association of stability of placement in foster care with cognitive, physical, and neural outcomes and symptoms of psychopathology at each assessment wave. β = association of age of placement, adjusting for covariates. β-values, 95% CIs, and *p*-values are bootstrapped (1,000 iterations).

	Assessment wave	β	95% CI	*p*
Cognitive, physical, and neural	30 months	-0.46	-0.92, 0.01	.064
	42 months	-0.17	-0.49, 0.16	.404
	54 months	-0.13	-0.63, 0.40	.616
	8 years	0.30	0.04, 0.54	.028
	12 years	0.21	-0.07, 0.47	.106
	16 years	0.33	0.06, 0.57	.002
Psychopathology	30 months	-0.22	-0.64, 0.17	.262
	42 months	<0.01	-0.34, 0.32	.992
	54 months	0.05	-0.19, 0.31	.713
	8 years	-0.12	-0.35, 0.10	.304
	12 years	-0.24	-0.41, -0.08	.004
	16 years	-0.22	-0.40, -0.40	.020

References

1. Nelson CA, Zeanah CH, Fox NA, et al: Cognitive recovery in socially deprived young children: the Bucharest Early Intervention Project. Science 2007; 318:1937–1940

2. Smyke AT, Koga SF, Johnson DE, et al: The caregiving context in institution-reared and family-reared infants and toddlers in Romania. J Child Psychol Psychiatry 2007; 48:210–218

3. Bayley N: Bayley Scales of Infant Development. 2nd ed. New York, NY, Psychological Corporation, 1993

4. Lindsey JC, Brouwers P: Intrapolation and extrapolation of age-equivalent scores for the Bayley II: a comparison of two methods of estimation. Clin Neuropharmacol 1999; 22:44–53

5. Marshall PJ, Reeb BC, Fox NA, et al: Effects of early intervention on EEG power and coherence in previously institutionalized children in Romania. Dev Psychopathol 2008; 20:861–880

6. Vanderwert RE, Marshall PJ, Nelson CA III, et al: Timing of intervention affects brain electrical activity in children exposed to severe psychosocial neglect. PLOS ONE 2010; 5:e11415

7. Debnath R, Tang A, Zeanah CH, et al: The long-term effects of institutional rearing, foster care intervention and disruptions in care on brain electrical activity in adolescence. Dev Sci 2020; 23:e12872

8. Vanderwert RE, Zeanah CH, Fox NA, et al: Normalization of EEG activity among previously institutionalized children placed into foster care: a 12-year follow-up of the Bucharest Early Intervention Project. Dev Cogn Neurosci 2016; 17:68–75

9. Sheridan MA, Fox NA, Zeanah CH, et al: Variation in neural development as a result of exposure to institutionalization early in childhood. Proc Natl Acad Sci U S A 2012; 109:12927–12932

10. Wade M, Fox NA, Zeanah CH, et al: Long-term effects of institutional rearing, foster care, and brain activity on memory and executive functioning. Proc Natl Acad Sci U S A 2019, 116:1808–1813

11. Debnath R, Buzzell GA, Morales S, et al: The Maryland analysis of developmental EEG (MADE) pipeline. Psychophysiology 2020; 57:13580

12. Delorme A, Makeig S: EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 2004; 134:9–21

13. Widmann A: Firfilt EEGLAB Plugin. Leipzig, University of Leipzig

14. Nolan H, Whelan R, Reilly RB: FASTER: fully automated statistical thresholding for EEG artifact rejection. J Neurosci Methods 2010; 192:152–162

15. Mognon A, Jovicich J, Bruzzone L, et al: ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology 2011; 48:40

16. Leach SC, Morales S, Bowers ME, et al: Adjusting ADJUST: optimizing the ADJUST algorithm for pediatric data using geodesic nets. Psychophysiology 2020; 57:13566

17. Machado S, Portella CE, Silva JG, et al: Changes in quantitative EEG absolute power during the task of catching an object in free fall. Arq Neuropsiquiatr 2007; 65:633–636

18. Marshall PJ, Fox NA, Group BC: A comparison of the electroencephalogram between institutionalized and community children in Romania. J Cogn Neurosci 2004; 16:1327–1338

19. Wechsler D: Wechsler Preschool and Primary Scale of Intelligence. San Antonio, TX, Harcourt Assessment, 2000

20. Wechsler D: The Wechsler Intelligence Scale for Children—Fourth Edition (WISC-IV). San Antonio, TX, The Psychological Corporation, 2003

21. Carter AS, Briggs-Gowan MJ, Jones SM, et al: The Infant-Toddler Social and Emotional Assessment (ITSEA): factor structure, reliability, and validity. J Abnorm Child Psychol 2003; 31:495–514

22. Smyke A, Zeanah CH: Disturbances of Attachment Interview. New Orleans, LA, Tulane University School of Medicine, 1999

23. Egger HL, Angold A: The preschool age psychiatric assessment (PAPA): a structured parent interview for diagnosing psychiatric disorders in preschool children, in Handbook of Infant, Toddler, and Preschool Mental Health Assessment. Edited by DelCarmen-Wiggens R, Carter A. New York, NY, Oxford University Press, 2004, pp 223–243

24. Essex MJ, Boyce WT, Goldstein LH, et al: The confluence of mental, physical, social, and academic difficulties in middle childhood. II: developing the MacArthur Health and Behavior Questionnaire. J Am Acad Child Adolesc Psychiatry 2002; 41:588–603

25. Almas AN, Degnan KA, Nelson CA, et al: IQ at age 12 following a history of institutional care: findings from the Bucharest Early Intervention Project. Dev Psychol 2016; 52:1858–1866

26. Humphreys KL, Gleason MM, Drury SS, et al: Effects of institutional rearing and foster care on psychopathology at age 12 years in Romania: follow-up of an open, randomised controlled trial. Lancet Psychiatry 2015; 2:625–634