Supplemental Materials

Analytic Sample

In **Supplemental Table 1**, we describe the distribution of study covariates by our analytic sample and others. Specifically, we compared the distributions of covariates between participants included in the total analytic sample (n=4438) and three other subsamples of ALSPAC: (1) participants who had at least one measure of social cognition, but were excluded based on other selection criteria (n=5239); (2) subset of the analytic sample who were exposed to sexual or physical abuse before age 10 (n=590); and (3) subset of the analytic sample who were exposed to caregiver physical or emotional abuse before age 10 (n=774).

Data Selection

Out of ALSPAC's 14,901 enrolled children alive at 1 year of age, there were 9677 children who had at least one measure of social cognition out of the three timepoints in which it was measured (ages 7.5, 10.5 and 14 years). With this base set of 9677 children, we then applied our exclusion criteria to identify the analytic sample. In a stepwise fashion, we first excluded children who did not have outcome measures at all three timepoints as would be needed for our life course modeling approach (5222 children remained). As the social experience of twins likely differs from singletons, we then excluded an additional 96 multiple-birth children. Lastly, as we restricted the analyses to only those children who had mothers and maternal figures as the sole reporters of their social cognition skills over the three timepoints of assessment to minimize reporter inconsistency, 688 children were additionally removed, yielding a total sample of 4,438 children.

We compared the distribution of covariates and outcome scores between the analytic sample (n=4438) and the subset of excluded participants with complete social cognition outcome data at all three time points, but who lacked consistent maternal reports (n=688). These two samples had largely similar sociodemographic characteristics and social cognition scores at all three time points; however, the excluded sample (without consistent maternal reports) were more likely to be born to mothers with slightly higher education (**Supplemental Table 2**).

Measures

ALSPAC was established to better understand how genetic and environmental features influence health and development of children (Fraser, 2013). Due to the breadth of this research question, specific scales or measures may not have been included at every time point of assessment. In **Supplemental Table 3**, we describe how each of the child maltreatment variables, described below was specially constructed including questions and time periods that were covered.

Child Maltreatment

To measure physical or sexual abuse by anyone, two items from an inventory assessing exposure to a series of life events were used. Specifically, mothers reported whether or not the child had been either "sexually abused" or "physically hurt by someone." If they answered "yes" on either of the two items, the child was coded as exposed. To measure caregiver physical or emotional abuse, both the mother and the partner provided responses to the following four items: 1) your partner was physically cruel to your children; 2) you were physically cruel to your children; 3) your partner was emotionally cruel to your children; 4) you were emotionally cruel to your children. If either the mother or the partner answered affirmatively to any of the four questions above, the child was coded as exposed. The participants were assured that their responses were confidential and no information would be reported to child welfare agencies, as no mandatory reporting laws were in place in the UK at the time of data collection (Bell, 1994; Khan, 2018). We note that because of the questionnaire wording both measures of child maltreatment ("Caregiver physical or emotional abuse" and "Sexual or physical abuse") could double-count caregiver physical abuse. The specific time periods covered by these questions are described in **Supplemental Table 3.**

Correlations between caregiver physical and emotional abuse items are shown in **Supplemental Table 4**. Correlations between the two types of maltreatment examined in this study are shown in **Supplemental Table 5**. Of note, while the prevalence of being ever exposed to sexual abuse before age 10 was much lower in the analytic sample (0.4%) compared to the prevalence of being ever exposed to physical abuse before age 10 (13.1%), the two exposures were moderately correlated ($r_{tetrachoric} = 0.39$).

Social Cognition

The distribution of social cognition scores across time, stratified by child sex, are shown in **Supplemental Table 6.**

Covariates

We controlled for the following covariates, measured at the time of the child's birth: *child race/ethnicity* (0=non-White; 1=White); *number of previous pregnancies* (between 0-3+); *maternal marital status* (0=never married; 1=widowed/divorced/separated; 2=married); *highest level of maternal education* (1=less than O-level, 2=O-level, 3=A-level, 4=Degree or above); *maternal age* (0=ages 15-19, 1=ages 20-35, 2=age>35); *homeownership* (0=mortgage/own home; 1=rent home; 2=other); *parent social class* (i.e. the highest social class of either parent: 1=professional; 2=managerial and technical; 3=skilled, non-manual; 4=skilled, manual; 5=semi-skilled, manual; 6=unskilled manual/other); and *maternal depressive symptoms* (measured by total scores on the Edinburgh Postnatal Depression Scale; scores ranged from 0-30 with higher scores indicating higher levels of depressive symptoms) (Adkins et al., 2011; Anney et al., 2010; Baker, Taylor & The Alspac Survey Team, 1997; Chen et al., 2013; Wood, White & Royston, 2008).

LARs Variable Selection and Structural Modeling

We achieved a single dataset for analysis by implementing LARs on the covariance structure among all variables, estimated by averaging the covariance structure across all multiply imputed datasets. This allowed us to avoid potential problems arising from different model selections across multiply imputed datasets (Wood et al., 2008).

We then evaluated the relative importance of these maltreatment variables using a two-stage structured lifecourse modeling approach (SLCMA) originally developed by Mishra (Mishra et al., 2009) for analyzing repeated, binary exposure data across the lifecourse. Relative to a more traditional regression model, the main advantage of the SLCMA is that it provides a structured and unbiased way to compare multiple competing theoretical models simultaneously and identify the most parsimonious explanation for the observed outcome variation.

In the first stage, we followed the approach of Smith (Smith et al., 2015) and entered the set of maltreatment variables described previously into a Least Angle Regression (LARs) procedure (Efron et al., 2004) in order to identify, separately for each type of maltreatment, the single theoretical model (or potentially more than one theoretical models working in combination) that explained the most variability in child social cognitive difficulties. We used a covariance test (Lockhart et al., 2014) and examined elbow plots (**Supplemental Figure 1**) to determine whether the selected models were supported by the ALSPAC data. Compared to other variable selection procedures, including stepwise regression, the SLCMA has been shown to not over-inflate effect size estimates (Efron et al., 2004) or bias hypothesis tests (Lockhart et al., 2014). Compared to other methods for the structured approach, LARs has been shown to have greater statistical power and not bias subsequent stages of analysis (Smith et al., 2015). To adjust for potential confounding, we regressed each encoded variable on the covariates and implemented LARs on the regression residuals (Smith et al., 2016).

In the second stage, the theoretical models determined by a covariance test p-value threshold of 0.05 in the first stage (which appeared before the elbow; see **Supplemental Figure 1**) was carried forward to a single multiple regression framework, where measures of effect would have been estimated for all selected hypotheses. The goal of this second stage was to determine the contribution of a selected theoretical model after adjustment for covariates as well as other selected theoretical models, in instances where more than one theoretical model was chosen in the first stage.

<u>Multiple Imputation</u>

As noted above, there were 4,438 children with complete outcome data at all three time points who met our inclusion criteria. However, a small proportion of these 4,438 children had missing exposure or covariate data; rates of missingness for exposure or covariate data ranged per variable from 4.3% (n=279 for maternal birth age) to 19.1% (n=1244 for presence versus absence of maternal psychopathology at 6 years).

To reduce potential bias and minimize loss of power due to attrition, we performed multiple imputation, separately for each exposure, using logistic regression in 20 datasets with 25 iterations each among all children with complete outcome data. In addition to imputing exposures, we also imputed covariates as described here. Of note, variables were included in the imputation models following the guidance of van Buuren and colleagues (van Buuren, Boshuizen & Knook, 1999; van Buuren & Groothuis-Oudshoorn, 2011) as well as prior research with imputation in the ALSPAC dataset (Evans et al., 2012; Ramchandani et al., 2008). The following variables were allowed to enter the imputation models: all covariates and exposures to the specific type of maltreatment from ages 0-8. Variables uncorrelated with the missing variable ($r<0.10$) were excluded from the imputation model (van Buuren et al., 1999; van Buuren & Groothuis-Oudshoorn, 2011). Imputation was performed with chained equations (Azur et al., 2011) with the *mice* package in R (van Buuren & Groothuis-Oudshoorn, 2011). To reduce noise in estimation of effect estimates, we did not impute the outcome (White, Royston & Wood, 2011). For each maltreatment, we assessed the convergence of the imputation model and the distribution of imputed data as compared to the observed data.

Results

Study results after winsorizing social cognition scores are shown in **Supplemental Table 7**.

<u>Exploring the Possibility that Social Cognition Predicts Child Maltreatment</u>

A primary hypothesis tested in this paper was that childhood maltreatment predicts future social cognitive skills. However, children with poor social cognitive skills may also be more likely than their peers to be exposed to child maltreatment. To explore this possibility, we performed a secondary analysis to examine the association between social cognition and child maltreatment. The first assessment of social cognition was available at age 7.5 years, which preceded the last two assessments of child maltreatment that we included in the analysis: sexual or physical abuse by anyone at 8 years and caregiver physical or emotional abuse at 9 years. We therefore fitted logistic regression models to test whether being abused later on (at 8 or 9 years) was predicted by levels of social cognition at 7.5 years. All baseline covariates included in our original analysis were also adjusted for here. Specifically, we assessed the associations between social cognition measured at age 7.5 years and odds of being exposed to each type of maltreatment separately in sex-stratified analyses (i.e., a total of four logistic regression models were fitted). We did not differentiate between incident cases of exposure to maltreatment at 8 or 9 years and cases with prior history of exposure, to preserve statistical power and keep the model parsimonious.

Among youth exposed to caregiver physical or emotional abuse at 9 years (n=158), there were 65 children whose parents had reported incident maltreatment, meaning children who had experienced new instances of caregiver physical or emotional abuse. Among youth exposed to physical or sexual abuse (by anyone) at 8 years (n=137), there were 59 were incident cases.

As shown in **Supplemental Table 8**, we found that poorer earlier social cognition skills were generally associated with lower levels of exposure to maltreatment. Specifically, the odds of being exposed to maltreatment were lower by 6-11% for each one-point increase on the social cognition scale (or worsening of social cognition scores). For example, for female participants, each one-point increase in social cognition at age 7.5 years was associated with a 9% decrease in the odds for being exposed to sexual or physical abuse by anyone at 8 years (OR=0.91, p=0.012). Similarly, each one-point increase in social cognition at age 7.5 was linked to a 11% decrease in the odds of being exposed to caregiver physical or emotional abuse at 9 years (OR=0.89, p=0.0001).

However, for boys, social cognition scores were only associated with sexual or physical abuse. Taken together, these findings do not suggest the possibility that children with poor social cognitive skills are at a substantially higher risk than their peers to be exposed to child maltreatment.

Supplemental Table 1. Comparisons of baseline sociodemographic characteristics in the total analytic sample versus among three subsamples of ALSPAC participants

	Total analytic sample (n=4438)		Excluded from the analytic sample (n=5239)			Exposed to sexual or physical abuse (n=590)		Exposed to physical or emotional abuse (n=774)	
	%	N	%	N	*p*-value	%	N	%	N
Gender					0.2				
Males	50.29	2232	51.63	2705		58.98	348	49.1	380
Females	49.71	2206	48.37	2534		41.02	242	50.9	394
Race					<0.01				
White	97.12	4186	94.69	4265		96.19	555	95.74	720
Non-White	2.88	124	5.31	239		3.81	22	4.26	32
Maternal Education					<0.01				
Less than O-level	18.2	794	29.96	1381		16.41	96	17.23	132
O-level	36.35	1586	34.34	1583		32.31	189	34.99	268
A-level	27.5	1200	22.99	1060		29.91	175	30.42	233
Degree or Above	17.95	783	12.71	586		21.37	125	17.36	133
Maternal Marital Status					<0.01				
Never Married	11.96	523	17.49	822		14.95	87	13.12	100
Widowed/Divorced/Separated	4.73	207	5.43	255		5.5	32	6.96	53
Married	83.31	3644	77.08	3622		79.55	463	79.92	609
Home Ownership					<0.01				
Mortgage/own home	85.63	3724	75.81	3539		81.83	473	80.4	607
Rent home	11.89	517	21.1	985		15.57	90	15.76	119
Other	2.48	108	3.08	144		2.6	15	3.84	29
Age of Mother at Child Birth					<0.01				
Ages 15-19	1.4	62	3.49	170		1.86	11	0.78	6
Ages 20-35	89.78	3978	89.1	4341		87.8	518	89.52	692
Age >35	8.82	391	7.41	361		10.34	61	9.7	75
Parental Social Class					<0.01				
Professional	15.4	683	10	524		15.25	90	16.02	124
Managerial and technical	39.39	1748	30.65	1606		42.54	251	40.44	313
Skilled, non-manual	21.43	951	19.18	1005		20.34	120	18.6	144

Skilled, manual	5.5	244	6.74	353		5.25	31	6.59	51
Semi-skilled, manual	1.35	60	2.02	106		1.02	6	1.42	11
Unskilled manual/other	16.94	752	31.4	1645		15.59	92	16.93	131
Number of previous pregnancies					<0.01				
0	47.7	2072	44.35	2063		47.47	272	38.76	293
1	36.07	1567	35.32	1643		34.21	196	39.15	296
2	12.22	531	14.6	679		13.61	78	16.93	128
3+	4.01	174	5.74	267		4.71	27	5.16	39

We compared the distributions of baseline characteristics between participants included in the total analytic sample (n=4438) and three other subsamples of ALSPAC: (1) participants who had at least one measure of social cognition, but were excluded based on other selection criteria (n=5239); (2) subset of the analytic sample who was exposed to sexual or physical abuse before age 10 (n=590); and (3) subset of the analytic sample who was exposed to caregiver physical or emotional abuse before age 10 (n=774). Notably, the original eligible sample (N=9677) consisted of all children that had at least one measure of social cognition. We restricted these analyses to singleton births with complete outcome data who had mothers and maternal figures as the sole reporters of their social cognition skills over the three timepoints of assessment.

p-values were determined from chi-squared tests, assessing whether the distributions of categorical covariates were different across samples. Values corresponding to education level are presented in rank order from lowest education level (less than O or Ordinary level) to Degree.

Supplemental Table 2. Distributions of covariates and social cognition scores in the analytic sample versus the sample of participants who were excluded due to having non-maternal reports

	Analytic sample (i.e., participants with maternal reports at all three time points) (n=4438)	Participants with complete but non-maternal reports of social cognition scores (n=688)	Compared to the total analytic sample
	N (%)	N (%)	*p*-value
Gender			0.535
Males	2232 (50.3)	364 (51.6)	
Females	2206 (49.7)	341 (48.4)	
Race			0.069
Non-White	124 (2.9)	29 (4.3)	
White	4186 (97.1)	653 (95.7)	
Maternal Education			<0.001
Less than O-level	794 (18.2)	139 (20.1)	
O-level	1586 (36.4)	186 (26.9)	
A-level	1200 (27.5)	189 (27.4)	
Degree or Above	783 (17.9)	177 (25.6)	
Maternal Marital Status			0.388
Never Married	523 (12.0)	74 (10.7)	
Widowed/Divorced/Separated	207 (4.7)	27 (3.9)	
Married	3644 (83.3)	588 (85.3)	
Home Ownership			0.312
Mortgage/own home	3724 (85.6)	581 (84.9)	
Rent home	517 (11.9)	91 (13.3)	
Other	108 (2.5)	12 (1.8)	
Age of Mother at child birth			0.808
Ages 15-19	62 (1.4)	8 (1.1)	
Ages 20-35	3978 (89.8)	630 (89.6)	
Age >35	391 (8.8)	65 (9.2)	
Parental Social Class			0.049
Professional	683 (15.4)	140 (19.9)	
Managerial and technical	1748 (39.4)	274 (38.9)	
Skilled, non-manual	951 (21.4)	127 (18.0)	
Skilled, manual	244 (5.5)	38 (5.4)	
Semi-skilled, manual	60 (1.4)	9 (1.3)	
Unskilled manual/other	752 (16.9)	117 (16.6)	

Number of previous pregnancies			0.729
0	2072 (47.7)	342 (49.6)	
1	1567 (36.1)	236 (34.3)	
2	531 (12.2)	81 (11.8)	
3+	174 (4.0)	30 (4.4)	
	Mean (SD)	Mean (SD)	*p*-value
Maternal depressive symptoms	5.03 (4.43)	4.95 (4.44)	0.688
Social cognition scores 7.5 years	2.69 (3.53)	2.54 (3.32)	0.321
Social cognition scores 10 years	14.19 (3.38)	13.98 (2.90)	0.121
Social cognition scores 14 years	14.43 (3.56)	14.46 (3.56)	0.866

We compared the distributions of baseline characteristics between participants included in the total analytic sample (n=4438) and a subset of excluded participants who had complete outcome data at all three time points but non-maternal reports (n=688).
p-values were determined from chi-squared tests and t-tests assessing the differences between the distributions of baseline covariates and social cognition skills in the two samples.

Supplemental Table 3. Summary of the two maltreatment measures and the time periods covered by each item

	Description	Time period covered
Sexual or Physical Abuse	Exposure to sexual or physical abuse was determined through an item asking the mother to indicate whether or not the child had been exposed to either sexual or physical abuse from anyone. This question was included at seven time-points: child ages 1.5, 2.5, 3.5, 4.75, 5.75, 6.75, and 8 years.	1.5y: 0.5-1.5y 2.5y: 1.5-2.5y 3.5y: 1-3.5y 4.75y: 3-4.74y 5.75y: 1.25-5.75y 6.75y: 5-6.75y 8.5y: 7-8y
Caregiver Physical or Emotional Abuse	Exposure to physical or emotional abuse was determined through mailed questionnaires administered separately to the mother and the mother's partner. Children were coded as having been exposed to physical or emotional abuse if the mother, partner, or both responded affirmatively to any of the following items assessed over seven time-points: (1) Your partner was physically cruel to your children; (2) You were physically cruel to your children; (3) Your partner was emotionally cruel to your children; (4) You were emotionally cruel to your children. The seven-time points were: 8 months, 1.75, 2.75, 4, 5, 6, and 9 years.	8m: birth to 8m 1.75y: 8m-1.75y 2.75y: 1.5-2.75y 4y: 2.5-4y 5y: 4-5y 6y: 5-6y 9y: 6-9y

Supplemental Table 4. Tetrachoric correlations between caregiver physical and emotional abuse items

		Parental physical abuse						
	Age	8 mo	1.75y	2.75y	4y	5y	6y	9y
Parental emotional abuse	8 mo	0.78	--	--	--	--	--	--
	1.75y	0.60	0.73	--	--	--	--	--
	2.75y	0.45	0.67	0.75	--	--	--	--
	4y	0.36	0.48	0.56	0.78	--	--	--
	5y	0.47	0.50	0.56	0.58	0.79	--	--
	6y	0.45	0.40	0.41	0.61	0.55	0.70	--
	9y	0.36	0.36	0.41	0.41	0.41	0.41	0.77

Tetrachoric correlation coefficients are presented in each cell to show the pairwise correlation between caregiver physical and emotional abuse at each time point. Notably, the two measures, when measured at the same tie point (see the diagonal), were strongly correlated (*rho* > 0.7).

Supplemental Table 5. Tetrachoric correlations between types of childhood maltreatment

Physical or emotional abuse (N=3677)							
Age	8 mo	1.75	2.75	4	5	6	9
8 mo	1	--	--	--	--	--	--
1.75	0.72	1	--	--	--	--	--
2.75	0.59	0.72	1	--	--	--	--
4	0.46	0.64	0.73	1	--	--	--
5	0.51	0.55	0.61	0.63	1	--	--
6	0.49	0.58	0.56	0.64	0.68	1	--
9	0.44	0.49	0.39	0.42	0.50	0.51	1

Sexual or physical abuse (by anyone) (N=3689)							
Age	1.5	2.5	3.5	4.75	5.75	6.75	8
1.5	1	--	--	--	--	--	--
2.5	0.5	1	--	--	--	--	--
3.5	0.36	0.39	1	--	--	--	--
4.75	0.33	0.44	0.44	1	--	--	--
5.75	0.4	0.43	0.47	0.52	1	--	--
6.75	0.3	0.4	0.35	0.46	0.63	1	--
8	0.46	0.37	0.39	0.44	0.54	0.59	1

Note. These results were generated using non-imputed datasets.

Supplemental Table 6. Social cognition scores across time

	Girls		Boys	
Age	Mean	S.D.	Mean	S.D.
7.5	2.28	3.04	3.09	3.93
10.5	1.83	2.84	2.45	3.79
14	2.19	3.19	2.56	3.86

Note. At each time period of measurement, there was a significant difference ($p<0.001$) between boys' and girls' scores

Supplemental Table 7. Results of the SLCMA for each measure of maltreatment on social cognition that were winsorized at the 90% percentile to address data skewness

	Sexual or physical abuse (by anyone)						
	Stage 1			Stage 2			
	Model(s) selected	R^2	P Value	β	S.E.	Lower CI	Upper CI
Girls (N=2206)							
Age 7.5	Accumulation	0.36 %	0.01	0.39	0.10	0.19	0.59
Age 10.5	Accumulation	0.46 %	<0.01	0.33	0.08	0.18	0.48
Age 14	Accumulation	0.78 %	<0.01	0.43	0.91	0.25	0.61
Boys (N=2232)							
Age 7.5	None						
Age 10.5	Accumulation	0.65 %	<0.01	0.28	0.06	0.16	0.39
Age 14	Accumulation	0.72 %	<0.01	0.35	0.07	0.21	0.49

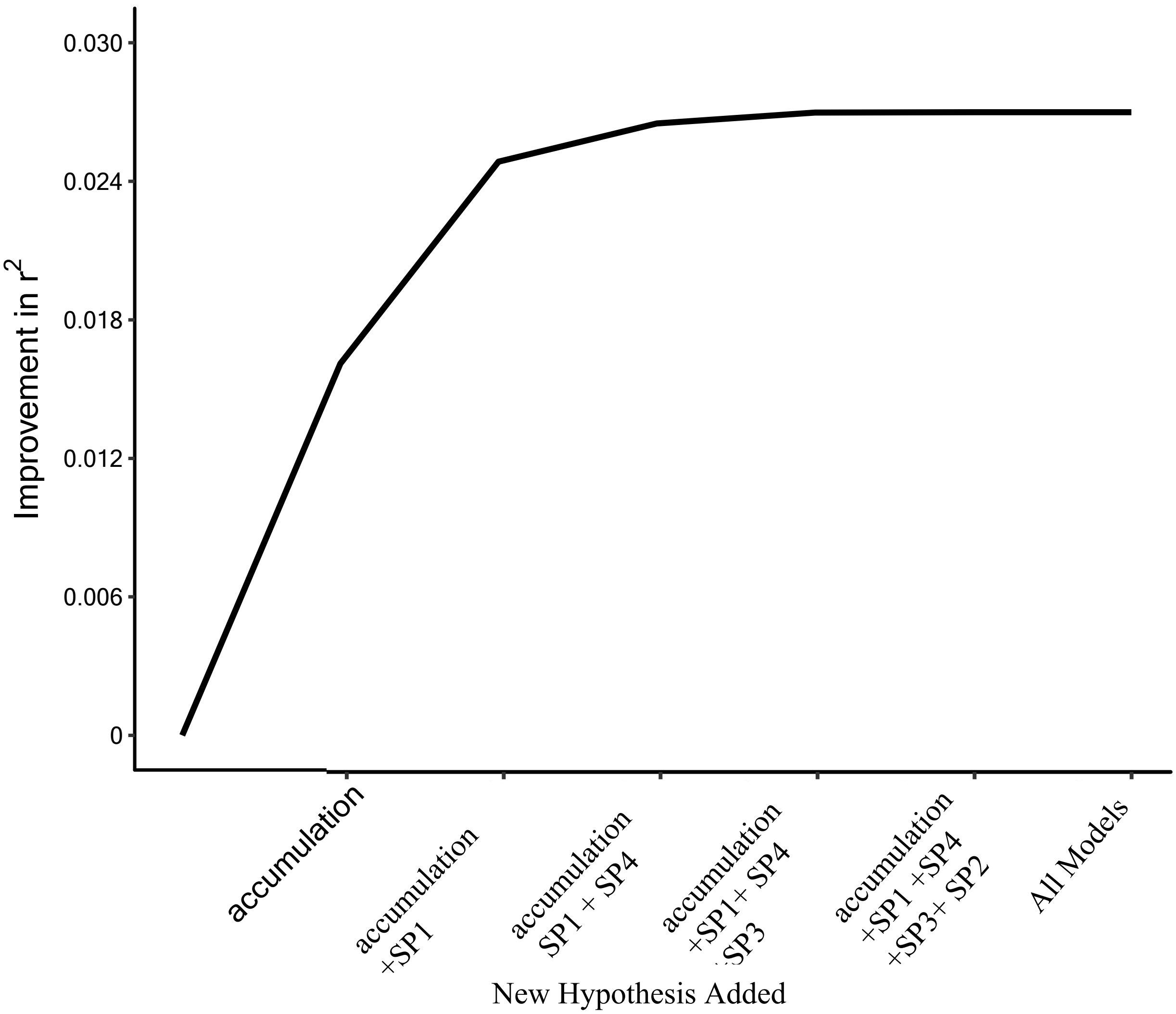
	Physical or emotional abuse						
	Stage 1			Stage 2			
	Model(s) selected	R^2	P Value	β	S.E.	Lower CI	Upper CI
Girls (N=2206)							
Age 7.5	None						
Age 10.5	None						
Age 14	None						
Boys (N=2232)							
Age 7.5	Ever Exposed	0.70 %	<0.01	0.58	0.14	0.31	0.85
Age 10.5	Accumulation	0.28 %	0.03	0.23	0.05	0.12	0.33
Age 14	None						

Stage 1 cell entries are r^2 values and p-values. Stage 2 cell entries are betas, standard errors, and p-values derived from multiple linear regression (one regression for each type of maltreatment) and social cognition measurement). Models were only reported at Stage 2 when the covariance test p-value was below the threshold of 0.1 .

Supplemental Table 8. Social cognition scores at age 7.5 predicting exposure to sexual or physical abuse at 8 years and caregiver physical or emotional abuse at 9 years

	Maltreatment Type	Beta	SE	P-value	OR	OR Lower CI	OR Upper CI
Female	Sexual or physical abuse 8 years	-0.10	0.04	0.0119	0.91	0.84	0.98
	Caregiver physical or emotional abuse 9 years	-0.12	0.03	0.0001	0.89	0.84	0.94
Male	Sexual or physical abuse 8 years	-0.06	0.02	0.0071	0.94	0.90	0.98
	Caregiver physical or emotional abuse 9 years	-0.01	0.03	0.7883	0.99	0.94	1.05

Supplemental Figure 1. Example elbow plot illustrating LARs variable selection procedure testing two life course models: accumulation and sensitive periods



LARs begins by first identifying the single variable with the strongest association to the outcome; it then identifies the combination of two variables with the strongest association, followed by three variables, and so on, until all variables are included. LARs therefore achieves parsimony by identifying the smallest combination of encoded variables that explain the most amount of outcome variation. In addition to a covariance test, which is calculated at each stage of the LARs procedure and tests the null hypothesis that adding the next encoded variable does not improve r^2, results can also be summarized in an "elbow plot," showing the increase in overall model r^2 as additional predictors are added to the model. The point where this plot levels off indicates the point of diminishing marginal improvement to the model goodness-of-fit from adding additional predictors, suggesting that the predictors included in the model at this point represent an optimal balance of parsimony and thoroughness. In this example, both accumulation and sensitive period 1 were selected in the best fitting models. SP =Sensitive Period.

References

Adkins, D. E., Aberg, K., McClay, J. L., Bukszar, J., Zhao, Z., Jia, P., Stroup, T. S., Perkins, D., McEvoy, J. P., Lieberman, J. A., Sullivan, P. F. & van den Oord, E. J. (2011). Genomewide pharmacogenomic study of metabolic side effects to antipsychotic drugs. *Mol Psychiatry 16*(3), 321-32.

Anney, R., Klei, L., Pinto, D., Regan, R., Conroy, J., Magalhaes, T. R., Correia, C., Abrahams, B. S., Sykes, N., Pagnamenta, A. T., Almeida, J., Bacchelli, E., Bailey, A. J., Baird, G., Battaglia, A., Berney, T., Bolshakova, N., Bolte, S., Bolton, P. F., Bourgeron, T., Brennan, S., Brian, J., Carson, A. R., Casallo, G., Casey, J., Chu, S. H., Cochrane, L., Corsello, C., Crawford, E. L., Crossett, A., Dawson, G., de Jonge, M., Delorme, R., Drmic, I., Duketis, E., Duque, F., Estes, A., Farrar, P., Fernandez, B. A., Folstein, S. E., Fombonne, E., Freitag, C. M., Gilbert, J., Gillberg, C., Glessner, J. T., Goldberg, J., Green, J., Guter, S. J., Hakonarson, H., Heron, E. A., Hill, M., Holt, R., Howe, J. L., Hughes, G., Hus, V., Igliozzi, R., Kim, C., Klauck, S. M., Kolevzon, A., Korvatska, O., Kustanovich, V., Lajonchere, C. M., Lamb, J. A., Laskawiec, M., Leboyer, M., Le Couteur, A., Leventhal, B. L., Lionel, A. C., Liu, X. Q., Lord, C., Lotspeich, L., Lund, S. C., Maestrini, E., Mahoney, W., Mantoulan, C., Marshall, C. R., McConachie, H., McDougle, C. J., McGrath, J., McMahon, W. M., Melhem, N. M., Merikangas, A., Migita, O., Minshew, N. J., Mirza, G. K., Munson, J., Nelson, S. F., Noakes, C., Noor, A., Nygren, G., Oliveira, G., Papanikolaou, K., Parr, J. R., Parrini, B., Paton, T., Pickles, A., Piven, J., Posey, D. J., Poustka, A., Poustka, F., Prasad, A., Ragoussis, J., Renshaw, K., Rickaby, J., Roberts, W., Roeder, K., Roge, B., Rutter, M. L., Bierut, L. J., Rice, J. P., Salt, J., Sansom, K., Sato, D., Segurado, R., Senman, L., Shah, N., Sheffield, V. C., Soorya, L., Sousa, I., Stoppioni, V., Strawbridge, C., Tancredi, R., Tansey, K., Thiruvahindrapduram, B., Thompson, A. P., Thomson, S., Tryfon, A., Tsiantis, J., Van Engeland, H., Vincent, J. B., Volkmar, F., Wallace, S., Wang, K., Wang, Z., Wassink, T. H., Wing, K., Wittemeyer, K., Wood, S., Yaspan, B. L., Zurawiecki, D., Zwaigenbaum, L., Betancur, C., Buxbaum, J. D., Cantor, R. M., Cook, E. H., Coon, H., Cuccaro, M. L., Gallagher, L., Geschwind, D. H., Gill, M., Haines, J. L., Miller, J., Monaco, A. P., Nurnberger, J. I., Jr., Paterson, A. D., Pericak-Vance, M. A., Schellenberg, G. D., Scherer, S. W., Sutcliffe, J. S., Szatmari, P., Vicente, A. M., Vieland, V. J., Wijsman, E. M., Devlin, B., Ennis, S. & Hallmayer, J. (2010). A genome-wide scan for common alleles affecting risk for autism. *Hum Mol Genet 19*(20), 4072-82.

Azur, M. J., Stuart, E. A., Frangakis, C. & Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res 20*(1), 40-9.

Baker, D., Taylor, H. & The Alspac Survey Team, H. (1997). The relationship between condition-specific morbidity, social support and material deprivation in pregnancy and early motherhood. *Social Science & Medicine 45*(9), 1325-36.

Chen, Y.-a., Lemire, M., Choufani, S., Butcher, D. T., Grafodatskaya, D., Zanke, B. W., Gallinger, S., Hudson, T. J. & Weksberg, R. (2013). Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics 8*(2), 203-9.

Efron, B., Hastie, T., Johnstone, I. & Tibshirani, R. (2004). Least angle regression. *The Annals of Statistics 32*(2), 407-99.

Evans, J., Melotti, R., Heron, J., Ramchandani, P., Wiles, N., Murray, L. & Stein, A. (2012). The timing of maternal depressive symptoms and child cognitive development: a longitudinal study. *Journal of child psychology and psychiatry 53*(6), 632-40.

Lockhart, R., Taylor, J., Tibshirani, R. J. & Tibshirani, R. (2014). A significance test for the LASSO. *Ann Stat 42*(2), 413-68.

Mishra, G., Nitsch, D., Black, S., De Stavola, B., Kuh, D. & Hardy, R. (2009). A structured approach to modelling the effects of binary exposure variables over the life course. *International Journal of Epidemiology 38*(2), 528-37.

Ramchandani, P. G., O'Connor, T. G., Evans, J., Heron, J., Murray, L. & Stein, A. (2008). The effects of pre- and postnatal depression in fathers: a natural experiment comparing the effects of exposure to depression on offspring. *J Child Psychol Psychiatry 49*(10), 1069-78.

Smith, A. D., Hardy, R., Heron, J., Joinson, C. J., Lawlor, D. A., Macdonald-Wallis, C. & Tilling, K. (2016). A structured approach to hypotheses involving continuous exposures over the life course. *Int J Epidemiol*.

Smith, A. D., Heron, J., Mishra, G., Gilthorpe, M. S., Ben-Shlomo, Y. & Tilling, K. (2015). Model Selection of the Effect of Binary Exposures over the Life Course. *Epidemiology 26*(5), 719-26.

van Buuren, S., Boshuizen, H. C. & Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med 18*(6), 681-94.

van Buuren, S. & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software 45*, urn:issn:1548-7660.

White, I. R., Royston, P. & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med 30*(4), 377-99.

Wood, A. M., White, I. R. & Royston, P. (2008). How should variable selection be performed with multiply imputed data? *Stat Med 27*(17), 3227-46.