Supporting Information

The relation between infant freezing and the development of internalizing symptoms in adolescence: A prospective longitudinal study

Hannah C. M. Niermann*, Anna Tyborowska, Antonius H. N. Cillessen, Marjolein M. van Donkelaar, Femke Lammertink, Megan R. Gunnar, Barbara Franke, Bernd Figner, Karin Roelofs

*** Correspondence:** Hannah C. M. Niermann: hcm.niermann@gmail.com

Appendix S1 – *5-HTTLPR* genotyping

Genetic analyses were carried out at the Department of Human Genetics of the Radboud University Medical Center (Nijmegen, Netherlands). Saliva samples were collected using Oragene kits (DNA Genotek, Kanata, Canada; 99.7% genotypic concordance with blood; Rylander-Rudqvist, Hakansson, Tybring, & Wolk, 2006), and genomic DNA was extracted as specified by the manufacturer.

The promoter region of the serotonin transporter gene (*SLC6A4*) contains a variable number of tandem repeats polymorphism (VNTR), *5-HTTLPR*, with two frequent alleles (short [S] and long [L] alleles), and a single nucleotide polymorphism (SNP; rs25531) resulting in an A/G substitution in the *5-HTTLPR*. In combination with the VNTR, only the long/A allele (L_A) is associated with higher expression of the gene, whereas the long/G allele (L_G) is comparable to the S-allele, with lower levels of mRNA (Hu et al., 2006). Rs25531 was genotyped using Taqman (Applied Biosystems). The *5-HTTLPR* VNTR was genotyped using standard PCR protocols. After the PCR, fragment length analysis was performed on the ABI Prism 3730 Genetic Analyser (Applied Biosystems, Nieuwekerk a/d Ijssel, Netherlands), and results were analyzed with GeneMapper® Software, version 4.0 (Applied Biosystems). We classified participants based on expression level: high (L_A/L_A; n = 27), intermediate

(L_A/S, L_A/L_G; n = 47), and low (L_G/L_G, S/L_G, S/S; n = 21). No deviations from Hardy-Weinberg Equilibrium were found for the *5-HTTLPR* genotype at p <.05. Additionally, there was no significant difference in genotype frequencies between gender (*Estimate* = 0.41(0.24), z = 1.71, p = .088).

Appendix S2 – Socioeconomic status (SES) and attrition analysis

SES was determined for the 15-months assessment wave based on the level of education (along a 7-point scale) and level of occupation (along a 7-point scale) for both parents. To derive a single SES score per participant, we first standardized the levels of education and occupation separately per parent, summed those scores for each parent, and calculated an average score across parents. For single parents (n = 6), the level of education and occupation of the primary caregiver was used (see also Smeekens, Riksen-Walraven, & van Bakel, 2007). SES of the participants included in current analyses varied between -3.12 and 2.63 (M = -0.004, SD = 1.65).

Participants missing all self-reports of internalizing data (n = 9) did not differ in their freezing response (Y_t = 1.18, 95% CI [-0.03, 0.11]) nor in their SES score (Y_t = 2.67, 95% CI [-0.21, 4.56]) from participants included in the self-reported internalizing analyses.

Appendix S3 – Infant freezing to robot

During the robot-confrontation, the primary caregiver was present throughout the procedure, but instructed to remain uninvolved unless the infant became upset and needed assistance. The experimenter controlled the robot remotely by turning its lights and sounds on and off, and moving it forwards and backwards. The distance between the robot and the infant varied depending on the infant's interaction with the robot. The robot situation lasted on average 3 min (M = 186.97 sec; SD = 52.76).

The observed intra-class correlation for infant freezing based on 22% of the videotapes was .76, 95% CI [.66, .83] (as reported in the main text), which was similar to previous freezing observation reliability measures (Buss, Davidson, Kalin, & Goldsmith, 2004). Of our analysis sample, 98 of the infants showed at least one freezing episode (≥3 seconds; a freezing episode lasted on average 6.64 sec; SD = 4.09; range: 3-24 sec). In total, these 98 infants showed 355 freezing episodes (M = 3.06, SD = 2.36, range: 0-10), of which 11% were followed by approach, 18% by avoidance, and 8% by distress behavior, whereas the remaining 63% showed no clear pattern of behavior after freezing.

To verify whether freezing can be considered a partially separate construct that at the same time is associated with other closely related concepts, we correlated freezing in response to the robot confrontation with other temperamental fearfulness assessments taken during the same robot confrontation (i.e., referencing to primary caregiver, physical contact with primary caregiver, physical contact with robot, laughing at primary caregiver, laughing at robot, and crying; as determined previously by van Bakel and Riksen-Walraven (2004)). Infant freezing was positively associated with physical contact with primary caregiver, and negatively associated with physical contact with robot, laughing at primary caregiver, and laughing at robot. It was not associated with referencing to primary caregiver and crying. Additionally, infants' level of behavioral inhibition was positively correlated with infants' level of freezing. Behavioral inhibition was taken as a single summary score of (i) latency to the infant's first vocalization, (ii) latency to touch the toy, and (iii) duration of proximity to primary caregiver determined in both the robot and stranger conditions (Fox, Henderson, Rubin, Calkins, & Schmidt, 2001). However, infants' freezing was not associated with social fearfulness, assessed via parent-reported Toddler Behavior Assessment Questionnaire at 15-months of age (TBAQ; Goldsmith & Rothbart, 1991). For statistical details of these correlational results, see Table S1. These correlational analyses suggest that freezing is correlated with other temperamental fearfulness measures, but at the same time these relatively low correlations (i.e., Kendall's tau correlation coefficient) suggest that freezing can be considered a partially separate construct.

Table S1

Correlational results (with 95% confidence intervals in parentheses) of infant freezing behavior to robot confrontation with other temperamental fearfulness assessments.

	Infant freezing
Physical contact with primary caregiver	.17 [.03, .31]*
Physical contact with robot	-.24 [-.38, -.10]**
Laughing at primary caregiver	-.17 [-.27, -.06]*
Laughing at robot	-.19 [-.33, -.06]**
Referencing to primary caregiver	-.10 [-.23, .05]
Crying	-.09 [-.23, .05]
Behavioral inhibition	.16 [.04, .28]**
Social fearfulness	-.04 [-.16, .06]

Notes: * $p < .05$, ** $p < .01$, We used the rank based correlation coefficient Kendall's tau because of the constrained (i.e., between 0 and 1) distribution of infant freezing, but also of the temperamental fearfulness assessments (i.e., rated as either 1 [= not at all present], 2 [= sometimes present], or 3 [= often present]) taken during this robot confrontation, determined previously by van Bakel and Riksen-Walraven (2004).

Appendix S4 – Infant freezing to stranger

We also assessed infant freezing during a stranger situation. During this situation, a woman unfamiliar to the infant entered the room with a toy ladybug containing colorful blocks and sat quietly for approximately 1 min. The woman then started to play with the ladybird and the blocks and invited the infant to play with them as well (~2 min; van Bakel & Riksen-Walraven, 2004). Similar to infant freezing in the robot context, a trained coder scored all the videos. Reliability was determined using 22% of the videotapes that were joint-coded with another trained coder, resulting in an intra-class correlation of .96, 95% CI [.93, .98]. Proportion scores were computed to control for different lengths in each stranger situation (*M*

= 182.58 sec, *SD* = 35.02). Only 34 of the 116 infants showed freezing episodes (one or more) in this situation (*M* = 0.01, *SD* = 0.03, range: 0.00-0.19).

Infant freezing in response to the stranger context did not predict relative changes in self- and parent-reported internalizing symptoms (self-reported internalizing symptoms: linear age × infant freezing: $\chi^2(1) = 0.03$, $p > .250$, 95% CI [-0.02, 0.02]; quadratic age × infant freezing: $\chi^2(1) = 0.83$, $p > .250$, 95% CI -0.02, 0.01]; parent-reported internalizing symptoms: linear age × infant freezing: $\chi^2(1) = 0.17$, $p > .250$, 95% CI [-0.003, 0.004]; quadratic age × infant freezing: $\chi^2(1) = 0.28$, $p > .250$, 95% CI [-0.002, 0.003]).[1] However, we observed a linear age effect for parent-reported internalizing symptoms, $\chi^2(1) = 6.92$, $p = .010$, 95% CI [0.001, 0.01], suggesting that participants showed an overall increase in parent-reported internalizing symptoms during development. Similar non-significant results were observed when gender was included as a main effect as well as in interaction with infant freezing and the linear and quadratic effects of age (self-reported internalizing symptoms: linear age × infant freezing: $\chi^2(1) = 0.44$, $p > .250$, 95% CI [-0.03, 0.01]; quadratic age × infant freezing: $\chi^2(1) = 0.10$, $p > .250$, 95% CI [-0.02, 0.02]; parent-reported internalizing symptoms: linear age × infant freezing: $\chi^2(1) = 0.34$, $p > .250$, 95% CI [-0.01, 0.003]; quadratic age × infant freezing: $\chi^2(1) = 0.00$, $p > .250$, 95% CI [-0.003, 0.003]). However, we could replicate the previously observed linear age × gender interaction (self-reported internalizing symptoms: $\chi^2(1) = 11.14$, $p = .002$, 95% CI [0.01, 0.05]; parent-reported internalizing symptoms: $\chi^2(1) = 5.96$, $p = .022$, 95% CI [0.001, 0.01]). When including parental and peer stress as main effects

[1] As only 34 of our participants showed infant freezing in response to the stranger context, we repeated the analyses but this time including infant freezing to the stranger as a categorical variable (freezing vs no freezing). These models also suggest that the effect of infant freezing (stranger context) on internalizing symptoms was not significant.

as well as in interaction with infant freezing and the linear and quadratic effects of age, similar non-significant associations were observed. See for results of the full models Table S2 and S3.

Table S2

Test statistics predicting self-reported internalizing symptoms from linear effect of age, quadratic effect of age, infant freezing to stranger, and their interactions (Model 1). Model 2 adds gender as main and interaction effects. Model 3 adds parental (Parent) and peer stress (Peer) as main and interaction effects.

	Model 1: Infant freezing			**Model 2: Gender**			**Model 3: Social stress**		
Effect	χ^2**(df = 1)**	***p***	**95% CI**	χ^2**(df = 1)**	***p***	**95% CI**	χ^2**(df = 1)**	***p***	**95% CI**
Linear age	0.02	> .250	-0.02, 0.02	0.04	> .250	-0.01, 0.02	0.19	> .250	-0.01, 0.02
Quadratic age	0.16	> .250	-0.02, 0.01	0.09	> .250	-0.02, 0.01	0.15	> .250	-0.02, 0.01
Freezing	0.06	> .250	-0.001, 0.001	0.12	> .250	-0.19, 0.13	0.01	> .250	-0.15, 0.13
Linear age × Freezing	0.03	> .250	-0.02, 0.02	0.44	> .250	-0.03, 0.01	0.30	> .250	-0.01, 0.03
Quadratic age × Freezing	0.83	> .250	-0.02, 0.01	0.10	> .250	-0.02, 0.02	0.13	> .250	-0.02, 0.01
Gender				12.87	.003**	0.11, 0.37			
Linear age × Gender				11.14	.002**	0.01, 0.05			
Quadratic age × Gender				0.35	> .250	-0.02, 0.01			
Freezing × Gender				0.03	> .250	-0.17, 0.15			
Linear age × Freezing × Gender				0.15	> .250	-0.02, 0.02			
Quadratic age × Freezing × Gender				0.11	> .250	-0.02, 0.02			
Parent							2.61	.117	-0.24, 0.02
Peer							1.07	> .250	-0.19, 0.06
Linear age × Parent							1.23	> .250	-0.01, 0.03
Linear age × Peer							0.38	> .250	-0.02, 0.01
Quadratic age × Parent							0.92	> .250	-0.02, 0.01
Quadratic age × Peer							0.05	> .250	-0.01, 0.02
Freezing × Parent							3.48	.065	-0.01, 0.31
Freezing × Peer							2.10	.166	-0.19, 0.03
Linear age × Freezing × Parent							0.86	> .250	-0.04, 0.01
Linear age × Freezing × Peer							0.01	> .250	-0.02, 0.02
Quadratic age × Freezing × Parent							1.53	.226	-0.03, 0.01
Quadratic age × Freezing × Peer							1.51	.219	-0.005, 0.02

Table S3

Test statistics predicting parent-reported internalizing symptoms from linear effect of age, quadratic effect of age, infant freezing to stranger, and their interactions (Model 1). Model 2 adds gender as main and interaction effects. Model 3 adds parental (Parent) and peer stress (Peer) as main and interaction effects.

	Model 1: Infant freezing			**Model 2: Gender**			**Model 3: Social stress**		
Effect	χ^2**(df = 1)**	***p***	**95% CI**	χ^2**(df = 1)**	***p***	**95% CI**	χ^2**(df = 1)**	***p***	**95% CI**
Linear age	6.92	.010*	0.001, 0.01	6.46	.015*	0.001, 0.01	4.87	.031*	0.0004, 0.01
Quadratic age	1.34	.250	-0.004, 0.001	1.20	> .250	-0.004, 0.001	1.99	.146	-0.004, 0.001
Freezing	0.19	> .250	-0.03, 0.02	0.16	> .250	-0.04, 0.02	0.44	> .250	-0.03, 0.02
Linear age × Freezing	0.17	> .250	-0.003, 0.004	0.34	> .250	-0.01, 0.003	0.04	> .250	-0.004, 0.003
Quadratic age × Freezing	0.28	> .250	-0.002, 0.003	0.00	> .250	-0.003, 0.003	0.36	> .250	-0.002, 0.004
Gender				0.56	> .250	-0.01, 0.03			
Linear age × Gender				5.96	.022*	0.001, 0.01			
Quadratic age × Gender				1.71	.166	-0.001, 0.004			
Freezing × Gender				0.02	> .250	-0.03, 0.03			
Linear age × Freezing × Gender				0.81	> .250	-0.002, 0.01			
Quadratic age × Freezing × Gender				0.10	> .250	-0.003, 0.004			
Parent							0.55	> .250	-0.03, 0.01
Peer							3.98	.059	-0.05, -0.0004
Linear age × Parent							5.20	.030*	0.001, 0.01
Linear age × Peer							3.18	.085	-0.01, 0.0003
Quadratic age × Parent							0.68	> .250	-0.002, 0.004
Quadratic age × Peer							0.09	> .250	-0.003, 0.002
Freezing × Parent							1.21	> .250	-0.01, 0.04
Freezing × Peer							1.32	> .250	-0.01, 0.03
Linear age × Freezing × Parent							0.97	> .250	-0.002, 0.01
Linear age × Freezing × Peer							2.74	.105	-0.0005, 0.01
Quadratic age × Freezing × Parent							0.11	> .250	-0.003, 0.004
Quadratic age × Freezing × Peer							0.77	> .250	-0.001, 0.003

Appendix S5 – Self-reported internalizing measures

Participants rated their own internalizing behavior using different anxiety and depression questionnaires at ages 9, 12, 14, and 17. These questionnaires reflect reliable, valid, and age-appropriate instruments to measure anxiety and depression. At age 9, we used the Short Depression Inventory for Children (SDIC; De Wit, 1987) and a shortened version of the Revised Children's Manifest Anxiety Scale (RCMAS; consisting of 6 items; Reynolds & Richmond, 1978; Reynolds & Richmond, 1985). The SDIC and RCMAS have good psychometric properties (Meijer, Mellenbergh, & de Wit, 1986; Reynolds & Richmond, 1978) and have been used previously in 9-year-olds to assess depression and anxiety respectively (Boer, Smit, Morren, Roorda, & Yzermans, 2009; Dadds, Spence, Holland, Barrett, & Laurens, 1997; Jansen, van de Looij-Jansen, de Wilde, & Brug, 2008; Kendall, 1994; Silverman et al., 1999; van de Looij-Jansen, Jansen, de Wilde, Donker, & Verhulst, 2011). At age 12, only a depression questionnaire was administered. We used the Children's Depression Inventory (CDI; Kovacs, 1981; translated by Timbremont & Braet, 2002), which is a widely used and one of the most scrutinized measures for self-reported depression in adolescence (Saylor, Finch, Spirito, & Bennett, 1984; Smucker, Craighead, Craighead, & Green, 1986). At age 14, we used the CDI and the Social Anxiety Scale for Adolescence (SAS; La Greca, 1998). The SAS has shown to be a psychometrically sound instrument to measure social anxiety in adolescence (Inderbitzen-Nolan & Walters, 2000; La Greca & Lopez, 1998). At age 17, we used the SAS and the anxiety, depression[2], and somatic complaints subscales of the Symptom Checklist-90-R (SCL-90-R; Arrindell & Ettema, 1975, 1986, 2005). These subscales of the SCL-90-R have been used previously to measure internalizing symptoms in adolescence (Ge, Conger, & Elder, 2001; Kim, Conger, Elder, &

[2]Due to technical problems, one item of the depression subscale of the Symptom Checklist 90-R (SCL-90-R) was missing for each participant.

Lorenz, 2003). See Table 1 in the main text for descriptive information of these anxiety and depression questionnaires. Anxiety and depression scores were positively correlated with each other at each age time point (*r*s between .41-.78, all *p*s < .001).

Appendix S6 – Parent-reported internalizing measures

We used age-appropriate versions of the CBCL, namely the CBCL for 4-18-year-olds at age 5, the CBCL for 6-18-year-olds at ages 9, 12, and 14, and the CBCL for 4-18-year-olds at age 17. At age 17, 8 of our participants received the CBCL for 6-18-year-olds. There is only a small difference between these two versions of the CBCL regarding the assessment of internalizing symptomatology (26 items are the same, whereas 5 items are different in the CBCL for 4-18-year-olds and 6 items in the CBCL for 6-18-year-olds). In general, the CBCL has been shown to have good reliability and validity (Achenbach, 1991).

Appendix S7 – Linear mixed-effect models - Statistical information

Based on participants' birthdate and their date of participation, participants' actual age was calculated for each measurement wave (see Table 1 in the main text for descriptive information regarding participants' age). In case the date of participation was missing for a participant, the mean age of the remaining participants was used as an estimate (n = 35). We determined the polynomial (linear and quadratic; using the *poly* function in *stats* package; R Core Team, 2016) effects of age. To increase likelihood of convergence of the linear mixed-effect models, we multiplied the linear and quadratic age effects by 100. The repeated-measure nature of the internalizing data was taken into account by including a per-participant random intercept and by modeling the linear as well as the quadratic effect of age not only as fixed effects but also as random slopes varying across participants (all possible random correlation terms were also included). This represents a "maximal" random effects structure as recommended by Barr, Levy, Scheepers, and Tily (2013) to avoid inflated Type-1 errors.

P-values were determined using bootstrapped likelihood ratio tests (requested samples: 1000), using the function *mixed* of the package *afex* (Singmann, Bolker, & Westfall, 2015). Confidence intervals (CI) were determined using profile-based CI using the function *profile* and *confint* of the package *lme4* (version 1.1.10; Bates, Maechler, Bolker, & Walker, 2015). To check whether missingness occurred completely at random in our two dependent variables, we used Little's MCAR test. Little's MCAR test suggested that our data were missing completely at random (self-reported internalizing: $\chi^2(25) = 36.11$, $p = .070$; parent-reported internalizing: $\chi^2(58) = 65.99$, $p = .220$). Finally, we investigated whether our longitudinal data did not violate the assumption of independent residuals, as assumed by the function *lmer* (*lme4* package; version 1.1.10; Bates et al., 2015). We visually inspected the residuals of each model and conducted Durban Watson tests to formally test whether there was any evidence for an autocorrelative structure in our residuals. All Durban Watson test statistics ranged between > 2 and < 3, suggesting no evidence for autocorrelation (Field, 2009).

Table S4

Test statistics predicting self-reported internalizing symptoms from linear effect of age, quadratic effect of age, infant freezing to robot, and their interactions (Model 1). Model 2 adds gender as main and interaction effects. Model 3 adds parental (Parent) and peer stress (Peer) as main and interaction effects.

	Model 1: Infant freezing			**Model 2: Gender**			**Model 3: Social stress**		
Effect	**χ^2(df = 1)**	***p***	**95% CI**	**χ^2(df = 1)**	***p***	**95% CI**	**χ^2(df = 1)**	***p***	**95% CI**
Linear age	0.00	> .250	-0.02, 0.02	0.01	> .250	-0.01, 0.02	0.10	> .250	-0.01, 0.02
Quadratic age	0.10	> .250	-0.02, 0.01	0.17	> .250	-0.02, 0.01	0.48	> .250	-0.02, 0.01
Freezing	0.02	> .250	-0.14, 0.12	0.03	> .250	-0.11, 0.14	0.34	> .250	-0.09, 0.17
Linear age × Freezing	7.41	.007**	-0.04, -0.01	7.10	.009**	-0.04, -0.01	9.12	.006**	-0.04, -0.01
Quadratic age × Freezing	8.38	.007**	-0.04, -0.01	9.19	.002**	-0.04, -0.01	10.11	.002**	-0.04, -0.01
Gender				13.57	.002**	0.11, 0.36			
Linear age × Gender				8.12	.002**	0.01, 0.04			
Quadratic age × Gender				0.51	> .250	-0.02, 0.01			
Freezing × Gender				0.11	> .250	-0.10, 0.15			
Linear age × Freezing × Gender				0.05	> .250	-0.02, 0.01			
Quadratic age × Freezing × Gender				1.35	.246	-0.02, 0.01			
Parent							2.98	.106	-0.25, 0.02
Peer							1.66	.234	-0.21, 0.04
Linear age × Parent							2.72	.104	-0.003, 0.03
Linear age × Peer							0.18	> .250	-0.02, 0.01
Quadratic age × Parent							0.06	> .250	-0.02, 0.01
Quadratic age × Peer							0.01	> .250	-0.01, 0.02
Freezing × Parent							0.07	> .250	-0.12, 0.16
Freezing × Peer							0.28	> .250	-0.16, 0.09
Linear age × Freezing × Parent							1.00	> .250	-0.03, 0.01
Linear age × Freezing × Peer							1.24	> .250	-0.03, 0.01
Quadratic age × Freezing × Parent							1.69	.213	-0.01, 0.03
Quadratic age × Freezing × Peer							5.78	.020*	0.003, 0.03

Appendix S8 – Parent-reported internalizing symptoms

Infant freezing did not predict relative changes in parent-reported internalizing symptoms (linear age × infant freezing: $\chi^2(1) = 0.58$, $p > .250$, 95% CI [-0.005, 0.002]; quadratic age × infant freezing: $\chi^2(1) = 0.09$, $p > .250$, 95% CI [-0.002, 0.003]). We observed only a main effect of linear age, $\chi^2(1) = 4.95$, $p = .026$, 95% CI [0.0005, 0.01], indicating an overall increase in internalizing symptoms across development. Similar results were observed when gender was included (see Appendix S8 and Table S5). To investigate whether social environment moderated the association between infant freezing and relative changes in internalizing symptoms, we added parental and peer stress. As before, we observed only a main effect of linear age, $\chi^2(1) = 5.12$, $p = .028$, 95% CI [0.001, 0.01], but no moderation by parental or peer stress. For results of the full models, see Table S5.

Table S5

Test statistics predicting parent-reported internalizing symptoms from linear effect of age, quadratic effect of age, infant freezing to robot, and their interactions (Model 1). Model 2 adds gender as main and interaction effects. Model 3 adds parental (Parent) and peer stress (Peer) as main and interaction effects.

	Model 1: Infant freezing			Model 2: Gender			Model 3: Social stress		
Effect	χ^2(df = 1)	*p*	95% CI	χ^2(df = 1)	*p*	95% CI	χ^2(df = 1)	*p*	95% CI
Linear age	4.95	.026*	0.0005, 0.01	5.79	.019*	0.001, 0.01	5.12	.028*	0.001, 0.01
Quadratic age	1.59	.154	-0.004, 0.001	1.28	.212	-0.004, 0.001	2.87	.092	-0.004, 0.0003
Freezing	1.53	.220	-0.04, 0.01	1.47	.245	-0.04, 0.01	0.54	> .250	-0.03, 0.01
Linear age × Freezing	0.58	> .250	-0.005, 0.002	0.33	> .250	-0.004, 0.002	0.48	> .250	-0.005, 0.002
Quadratic age × Freezing	0.09	> .250	-0.002, 0.003	0.19	> .250	-0.002, 0.003	0.05	> .250	-0.003, 0.002
Gender				0.28	> .250	-0.02, 0.03			
Linear age × Gender				5.58	.029*	0.001, 0.01			
Quadratic age × Gender				0.90	> .250	-0.001, 0.004			
Freezing × Gender				0.07	> .250	-0.03, 0.02			
Linear age × Freezing × Gender				0.09	> .250	-0.004, 0.003			
Quadratic age × Freezing × Gender				0.13	> .250	-0.002, 0.003			
Parent							1.09	> .250	-0.04, 0.01
Peer							4.53	.035*	-0.05, -0.002
Linear age × Parent							2.78	.117	-0.001, 0.01
Linear age × Peer							4.14	.064	-0.01, -0.0001
Quadratic age × Parent							0.45	> .250	-0.002, 0.003
Quadratic age × Peer							0.01	> .250	-0.002, 0.002
Freezing × Parent							2.81	.099	-0.05, 0.004
Freezing × Peer							0.05	> .250	-0.02, 0.03
Linear age × Freezing × Parent							0.53	> .250	-0.01, 0.002
Linear age × Freezing × Peer							0.24	> .250	-0.003, 0.004
Quadratic age × Freezing × Parent							2.92	.081	-0.0004, 0.01
Quadratic age × Freezing × Peer							0.05	> .250	-0.003, 0.002

Appendix S9 - Gender

Self-reported internalizing symptoms. Because internalizing symptoms can vary by gender (Bongers, Koot, van der Ende, & Verhulst, 2003), we examined whether the pattern of effects of infant freezing in the robot context on relative changes in self-reported internalizing symptoms was shared across gender. Therefore, we ran our main model predicting relative changes in self-reported internalizing symptoms from infant freezing again. However, this time we included gender (male/female; sum-to-zero contrast) as a main effect as well as in interaction with infant freezing and the linear and quadratic effects of age. The observed interaction effects of infant freezing on relative changes in internalizing symptoms remained when gender was included as a main and interaction effect (linear age × infant freezing: $\chi^2(1) = 7.10$, $p = .009$, 95% CI [-0.04, -0.01]; quadratic age × infant freezing: $\chi^2(1) = 9.19$, $p = .002$, 95% CI [-0.04, -0.01]). This suggested that the pattern of observed effects for infant freezing on the relative changes in self-reported internalizing symptoms was shared across gender. We also observed a linear age × gender interaction, $\chi^2(1) = 8.12$, $p = .002$, 95% CI [0.01, 0.04]: female relative to male participants showed increased levels of self-reported internalizing symptoms during adolescence (Figure S1; see also Table S4 for results of the full model).

Parent-reported internalizing symptoms. Similar non-significant associations between infant freezing and the relative changes in the development of parent-reported internalizing symptoms (linear age × infant freezing: $\chi^2(1) = 0.33$, $p > .250$, 95% CI [-0.004, 0.002]; quadratic age × infant freezing: $\chi^2(1) = 0.19$, $p > .250$, 95% CI [-0.002, 0.003]) were observed when gender was included as a main effect and in interaction with infant freezing and the linear and quadratic effects of age. Similar to the model without gender, we observed a main effect of linear age, $\chi^2(1) = 5.79$, $p = .019$, 95% CI [0.001, 0.01], but this time also an interaction effect of linear age × gender, $\chi^2(1) = 5.58$, $p = .029$, 95% CI [0.001, 0.01]. Similar as observed for self-reported internalizing symptoms, female compared to male participants showed an increase in parent-reported internalizing symptoms during development (Figure S1; see Table S5 for results of the full model).

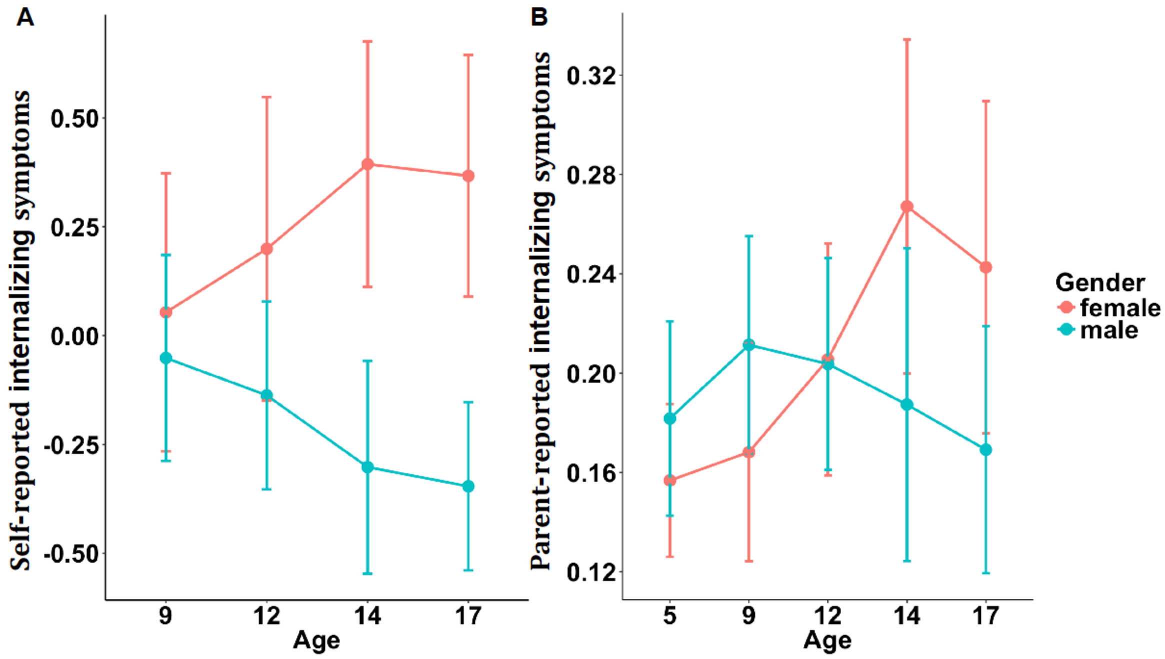

Figure S1. Panel A: Female relative to male participants showed an increase in self-reported internalizing symptoms from late childhood to late adolescence (age 9-17). Standardized scores were used for self-reported internalizing symptoms. Panel B: Similar as for self-reported internalizing symptoms, female participants showed an increase in parent-reported internalizing symptoms from mid childhood to late adolescence (age 5-17), while male participants showed an initial increase from age 5 to 9, followed by decrease in parent-reported internalizing symptoms from age 9 to 17. These panels show raw data of self- and parent-reported internalizing symptoms. Error bars represent 95% confidence intervals.

Appendix S10 – Externalizing symptoms

Because internalizing symptoms can occur in comorbidity with externalizing symptoms (i.e., stress-related and externally directed symptoms, including aggression and other forms of disruptive behavior) (Levy, Hawes, & Johns, 2014), we investigated whether externalizing symptoms could explain the observed association between infant freezing to robot and relative changes in self-reported internalizing symptoms by controlling for parent-reported externalizing symptoms as a time-invariant variable. Externalizing symptoms were added only as a main effect. Externalizing symptoms were assessed with the externalizing subscale of the parent-report Child Behavior Checklist (CBCL; Achenbach, 1991, 1992; Achenbach & Rescorla, 2001; α ranging from .87 to .89). Different versions of the CBCL were used (see Appendix S6). The mean scores of externalizing symptoms were positively correlated (*r*s ranging from .40 to .65, *p*s < .01) at ages 9, 12, 14, and 17. To determine a time-invariant score of externalizing symptoms across those ages, we standardized the mean scores of externalizing symptoms per age and averaged those across age. Missing data (24% missed one; 12% missed two; 3% missed three; 2% missed four measurement points) were handled by computing an adjusted score for participants with missing data such that only the non-missing observations were used to compute the average score.

To investigate potential externalizing effects, we added externalizing symptoms as a main effect to the described models in the main text. The observed effects of infant freezing on relative changes in self-reported internalizing symptoms as well as the observed moderation effect of peer stress were independent from the experience of externalizing symptoms (see Table S6 for results of these models).

Table S6

Test statistics predicting self-reported internalizing symptoms from linear effect of age, quadratic effect of age, infant freezing to robot, and their interactions (Model 1). Model 2 adds parental (Parent) and peer stress (Peer) as main and interaction effects. Please note that parent-reported externalizing symptoms is added as a time-invariant main effect to both models.

	Model 1: Infant freezing			**Model 2: Social stress**		
Effect	**χ^2 (df = 1)**	***p***	**95% CI**	**χ^2 (df = 1)**	***p***	**95% CI**
Linear age	0.06	> .250	-0.01, 0.02	0.12	> .250	-0.01, 0.02
Quadratic age	0.22	> .250	-0.02, 0.01	0.40	> .250	-0.02, 0.01
Freezing	0.12	> .250	-0.10, 0.15	0.44	> .250	-0.09, 0.18
Externalizing	4.55	.037*	0.01, 0.26	1.30	> .250	-0.06, 0.22
Linear age × Freezing	10.36	.003**	-0.04, -0.01	10.05	.002**	-0.04, -0.01
Quadratic age × Freezing	6.65	.008**	-0.03, -0.005	8.65	.007**	-0.04, -0.01
Parent				1.23	> .250	-0.22, 0.06
Peer				0.77	> .250	-0.19, 0.07
Linear age × Parent				1.78	.208	-0.01, 0.03
Linear age × Peer				0.17	> .250	-0.02, 0.01
Quadratic age × Parent				0.00	> .250	-0.02, 0.02
Quadratic age × Peer				0.00	> .250	-0.01, 0.01
Freezing × Parent				0.00	> .250	-0.15, 0.15
Freezing × Peer				0.36	> .250	-0.17, 0.09
Linear age × Freezing × Parent				0.10	> .250	-0.02, 0.02
Linear age × Freezing × Peer				1.02	> .250	-0.02, 0.01
Quadratic age × Freezing × Parent				0.51	> .250	-0.01, 0.02
Quadratic age × Freezing × Peer				5.59	.026*	0.003, 0.03

Appendix S11 – Self-reported anxiety and depression

Infant freezing. To explore whether the observed effects of infant freezing to the robot context on relative changes in self-reported internalizing symptoms were similar for self-reported depressive and anxiety symptoms, we repeated the linear mixed-effect model as reported in the main text, but this time predicting self-reported depressive and anxiety symptoms separately. Similar as observed for self-reported internalizing symptoms combined, we observed a linear age × infant freezing, $\chi^2(1) = 4.64$, $p = .031$, 95% CI [-0.04, -0.002], and a quadratic age × infant freezing interaction, $\chi^2(1) = 8.33$, $p = .007$, 95% CI [-0.04, -0.01], for self-reported depressive symptoms. This suggests that individuals who showed *longer* infant freezing behavior showed relatively higher levels of self-reported depressive symptoms at age 12, while after that time levels of depressive symptoms relatively decreased for these individuals (Figure S2C). At age 12, 6% of these individuals scored above a clinical cutoff score (Kovacs, 1981; Timbremont & Braet, 2002) on the CDI (compared to 4% of individuals with medium levels of infant freezing; no individuals with absent infant freezing scored above the clinical cutoff score). In contrast, individuals who showed *no* infant freezing displayed relatively higher levels of self-reported depressive symptoms at age 17 (Figure S2A; see Table S7). At age 17, 23% of these individuals scored in the *very high* range on the depression subscale of the SCL-90-R (Arrindell & Ettema, 2003; compared to 14% of individuals with medium levels of infant freezing; no individuals with longer infant freezing scored above the clinical cut-off score).

For self-reported anxiety symptoms, we only observed a linear age × infant freezing, $\chi^2(1) = 6.08$, $p = .017$, 95% CI [-0.04, -0.005], but not a quadratic age × infant freezing interaction, $\chi^2(1) = 1.05$, $p > .250$, 95% CI [-0.01, 0.02]: Infant freezing negatively predicted relative changes in self-reported anxiety symptoms (Figure S2D, Table S8). At age 17, 8% of individuals showing no infant freezing behavior scored in the *very high* range on the anxiety subscale of the SCL-90-R (Arrindell & Ettema, 2003; compared to 8% of individuals with longer levels of infant freezing and 10% of individuals with medium levels of infant freezing). Additionally, 20% of individuals with no freezing behavior scored above a clinical cut-off score on the SAS (La Greca, 1998; compared to 13% of individuals with longer infant freezing and 16% of individuals with medium levels of infant freezing).

The absence of a quadratic age × infant freezing interaction for self-reported anxiety symptoms might result from the missing anxiety measurement at age 12. To further explore whether the observed effects of infant freezing on self-reported depressive or anxiety symptoms were similar, we repeated the current linear mixed-effect model for depressive symptoms, but this time excluded depressive symptoms at age 12. Similar as for the model predicting anxiety symptoms (see Table S8), we were not able to include the random effect of the linear and quadratic slope of age simultaneously. Nevertheless, we observed similar results predicting self-reported depressive symptoms once with the linear slope of age (linear age × infant freezing: $\chi^2(1) = 2.82$, $p = .126$, 95% CI [-0.03, 0.003]; quadratic age × infant freezing: $\chi^2(1) = 1.21$, $p > .250$, 95% CI [-0.02, 0.01]) and once with the quadratic slope of age as random effect (linear age × infant freezing: $\chi^2(1) = 3.15$, $p = .084$, 95% CI [-0.03, 0.002]; quadratic age × infant freezing: $\chi^2(1) = 1.28$, $p > .250$, 95% CI [-0.03, 0.01]). These results suggest that the previously observed quadratic age × infant freezing interaction for self-reported depressive symptoms was driven by depressive symptoms at age 12. Similar as for self-reported anxiety symptoms, we observed a trend for a linear age × infant freezing interaction, suggesting that infant freezing negatively predicted changes in self-reported depressive symptoms.

Infant freezing and social environment. To explore whether the observed three-way-interaction between peer social preference × quadratic age × infant freezing on relative changes in self-reported internalizing symptoms was similar for self-reported depressive and anxiety symptoms, we repeated the linear mixed-effect model as described in the main text, but this time predicting self-reported depressive and anxiety symptoms separately. Similar as observed for self-reported internalizing symptoms, we observed a linear age × infant freezing, $\chi^2(1) = 3.90$, $p = .062$, 95% CI [-0.04, -0.0001], a quadratic age × infant freezing, $\chi^2(1) = 8.36$, $p = .003$, 95% CI [-0.04, -0.01], as well as a quadratic age × infant freezing × peer social preference interaction, $\chi^2(1) = 5.77$, $p = .013$, 95% CI [0.004, 0.04], for self-reported depressive symptoms (see Figures S2A-C and Table S7 in Supplementary Material). For self-reported anxiety symptoms, we also observed a linear age × infant freezing, $\chi^2(1) = 8.11$, $p = .004$, 95% CI [-0.04, -0.01], and a quadratic age × infant freezing × peer social preference interaction, $\chi^2(1) = 6.62$, $p = .009$, 95% CI [0.004, 0.03], but not a quadratic age × infant freezing interaction, $\chi^2(1) = 0.02$, $p > .250$, 95% CI [-0.02, 0.02] (Figures S2D-F, Table S8). Both depressive and anxiety symptoms contributed to the observed moderation effect of peer social preference on the association between infant freezing and relative changes in the development of internalizing symptoms.

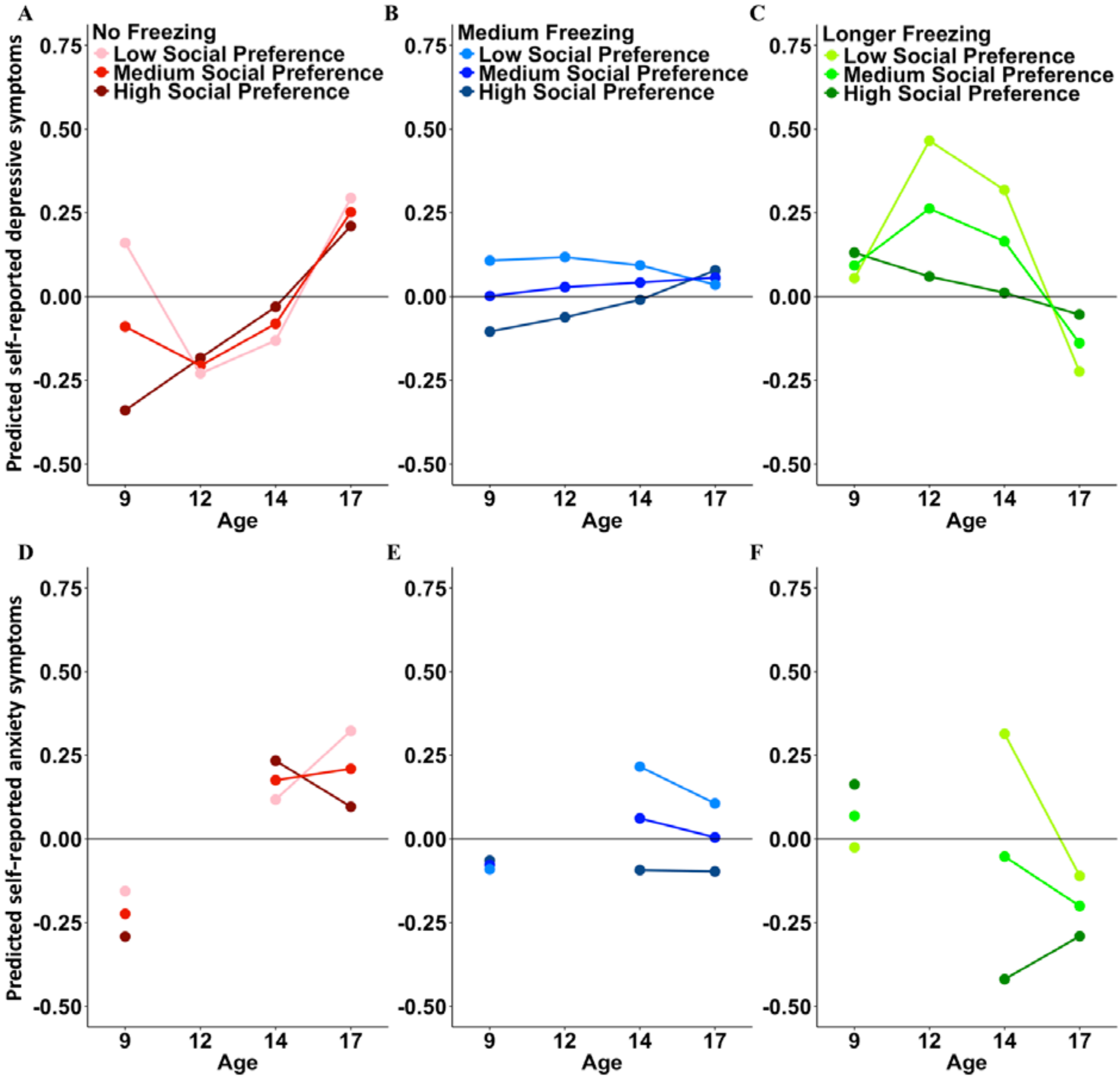

Figure S2. These line graphs show model-based changes in self-reported depressive (Panels A, B, C) and anxiety symptoms (Panels D, E, F) at ages 9, 12, 14, and 17 as a function of infant freezing and as a function of peer social preference (no freezing/low preference [1 *SD* below mean]; medium freezing/medium preference [$M = 0$]; longer freezing/high preference [1 *SD* above mean]). We administered only a depression, but not an anxiety questionnaire at age 12. Anxiety and depression self-reported symptoms were standardized per age.

Table S7

Test statistics predicting self-reported depressive symptoms from linear effect of age, quadratic effect of age, infant freezing to robot, and their interactions (Model 1). Model 2 adds parental (Parent) and peer stress (Peer) as main and interaction effects.

	Model 1: Infant freezing			Model 2: Social stress		
Effect	**χ^2 (df = 1)**	***p***	**95% CI**	**χ^2 (df = 1)**	***p***	**95% CI**
Linear age	0.00	> .250	-0.02, 0.02	0.20	> .250	-0.01, 0.02
Quadratic age	0.00	> .250	-0.02, 0.02	0.00	> .250	-0.02, 0.02
Freezing	0.02	> .250	-0.13, 0.14	0.99	> .250	-0.07, 0.20
Linear age × Freezing	4.64	.031*	-0.04, -0.002	3.90	.062	-0.04, -0.0001
Quadratic age × Freezing	8.33	.007**	-0.04, -0.01	8.36	.003**	-0.04, -0.01
Parent				4.73	.033*	-0.29, -0.02
Peer				0.75	> .250	-0.19, 0.08
Linear age × Parent				0.77	> .250	-0.01, 0.03
Linear age × Peer				0.84	> .250	-0.01, 0.03
Quadratic age × Parent				0.00	> .250	-0.02, 0.02
Quadratic age × Peer				0.20	> .250	-0.01, 0.02
Freezing × Parent				0.14	> .250	-0.18, 0.12
Freezing × Peer				0.00	> .250	-0.14, 0.13
Linear age × Freezing × Parent				3.18	.091	-0.04, 0.002
Linear age × Freezing × Peer				0.52	> .250	-0.03, 0.01
Quadratic age × Freezing × Parent[a]				0.06	> .250	-0.02, 0.02
Quadratic age × Freezing × Peer				5.77	.013*	0.004, 0.04

Notes: [a] We observed a convergence warning for the quadratic age × infant freezing × quality of parental behavior interaction for self-reported depressive symptoms. We carefully checked this warning by comparing the Log Likelihood estimates when using different optimizers for the model required to test the significance of the quadratic age × infant freezing × quality of parental behavior interaction term. As the different optimizers revealed very similar results, the original convergence warning can be treated as a false positive and the statistical results regarding the quadratic age × infant freezing × quality of parental behavior interaction can be assumed to be reliable.

Table S8

Test statistics predicting self-reported anxiety symptoms from linear effect of age, quadratic effect of age, infant freezing to robot, and their interactions (Model 1). Model 2 adds parental (Parent) and peer stress (Peer) as main and interaction effects.

	Model 1: Infant freezing			**Model 2: Social stress**		
Effect	**χ^2 (df = 1)**	***p***	**95% CI**	**χ^2 (df = 1)**	***p***	**95% CI**
Linear age	0.23	> .250	-0.01, 0.02	0.75	> .250	-0.01, 0.02
Quadratic age	0.01	> .250	-0.01, 0.01	0.69	> .250	-0.02, 0.01
Freezing	1.49	.231	-0.24, 0.06	0.47	> .250	-0.20, 0.09
Linear age × Freezing	6.08	.017*	-0.04, -0.005	8.11	.004**	-0.04, -0.01
Quadratic age × Freezing	1.05	> .250	-0.01, 0.02	0.02	> .250	-0.02, 0.02
Parent				0.24	> .250	-0.19, 0.11
Peer				1.02	> .250	-0.22, 0.07
Linear age × Parent				5.66	.017*	0.004, 0.04
Linear age × Peer				1.19	> .250	-0.03, 0.01
Quadratic age × Parent				0.86	> .250	-0.02, 0.01
Quadratic age × Peer				0.84	> .250	-0.01, 0.02
Freezing × Parent				0.67	> .250	-0.09, 0.22
Freezing × Peer				0.13	> .250	-0.17, 0.12
Linear age × Freezing × Parent				0.43	> .250	-0.02, 0.01
Linear age × Freezing × Peer				1.01	> .250	-0.03, 0.01
Quadratic age × Freezing × Parent				2.21	.168	-0.005, 0.03
Quadratic age × Freezing × Peer				6.62	.009**	0.004, 0.03

Notes: As we had no self-reported anxiety questionnaire at age 12, we were not able to include the random effect of the linear and quadratic slope of age at the same time into the model. Similar results were found when we conducted the model once with the linear slope of age and once with the quadratic slope of age as random effects. Here, the results of the linear slope of age as random effect are reported.

Appendix S12 - Serotonin Transporter Gene Polymorphism

As reported in the main text, we found that S'-homozygotes showed more deviations in both absent and longer infant freezing behavior, when compared to L'-carriers (see Figure S3).

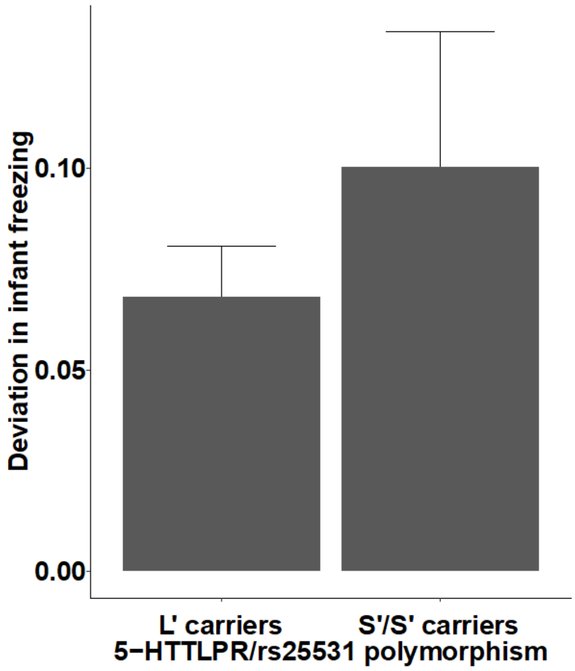

Figure S3. The bar graph illustrates that S'-homozygotes (S'/S' carriers) showed on average more deviations (in the sense of absent or longer infant freezing behavior) compared to L'-carriers (L'/L' and L'/S'-carriers combined) of the *5-HTTLPR/*rs25531 polymorphism. Error bars represent 95% confidence interval.

We also investigated whether the *5-HTTLPR/*rs25531 polymorphism moderates the association between deviant infant freezing (i.e., reflecting either absent or excessively long freezing; see main text for calculation of this deviant freezing score) and an individual's peak in self-reported internalizing symptoms. An individual's peak in internalizing symptoms was visually determined based on plotting the interactions between linear and quadratic age × infant freezing of the mixed-effect model results reported in the main text (model-based plotting as well as plotting of raw data yielded the same results): For individuals showing longer infant freezing (>1 *SD* above the freezing mean) we used the score of internalizing symptoms at age 12 as their peak in internalizing symptoms, for individuals showing no infant freezing (<1 *SD* below the freezing mean) we used their score of internalizing symptoms at age 17, whereas for individuals showing medium infant freezing (between >-1 *SD* and <1 *SD* of the freezing mean) we determined the average of internalizing symptoms between ages 12, 14, and 17 as they showed no clearly defined peak in internalizing symptoms. Using the *lm* function from the *stats* package (R Core Team, 2016), we ran a regression model[3] predicting peak in self-reported internalizing symptoms from deviant infant freezing, *5-HTTLPR/*rs25531 polymorphism (i.e., S'-homozygotes vs L'-carriers [L'/L' and L'/S' combined]), and their interaction (R^2 = .09, $F(3, 89) = 3.11$, $p = .030$; deviant freezing: $t(89) = 2.97$, $p = .004$; *5-HTTLPR/*rs25531: $t(89) = 0.76$, $p > .250$; deviant freezing × *5-HTTLPR/*rs25531: $t(89) = -1.65$, $p = .103$). These results suggest that *5-HTTLPR/*rs25531 did not moderate the association between deviant infant freezing and peak in internalizing symptoms, but that the effect of infant freezing on peak in internalizing symptoms remained when taking *5-HTTLPR/*rs25531 variations into account.

Appendix S13 – Association between infant freezing and peer stress

We observed no significant association between infant freezing to robot and peer stress for participants included in current self-reported internalizing analyses ($r = -.01$, $p > .250$, 95% CI [-.14, .10]; we used the rank based correlation coefficient Kendall's tau because of the constrained [i.e., between 0 and 1] distribution of infant freezing).

---

[3]To meet statistical assumptions, two influential cases (standardized outliers >3 *SD* combined with Cook's distance > 4/*N*) were removed.

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1992). *Manual for Child Behavior Checklist/ 2-3 and 1992 Profile*. Burlington, VT: University of Vermont, Dept. of Psychiatry.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Arrindell, W. A., & Ettema, J. H. M. (1975, 1986, 2005). *Symptom checklist: Handleiding bij een multidimensionale psychopathologie-indicator*. Amsterdam: Pearson Assessment and Information B.V.

Arrindell, W. A., & Ettema, J. H. M. (2003). *Handleiding bij een multidimensionele psychopathologie-indicator. Symptom Checklist SCL-90. (Herziene uitgave)*. Lisse: Swets & Zeitlinger.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang, 68*(3), 1-43. doi:10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01

Boer, F., Smit, C., Morren, M., Roorda, J., & Yzermans, J. (2009). Impact of a technological disaster on young children: A five-year postdisaster multiinformant study. *J Trauma Stress, 22*(6), 516-524. doi:10.1002/jts.20461

Bongers, I. L., Koot, H. M., van der Ende, J., & Verhulst, F. C. (2003). The normative development of child and adolescent problem behavior. *J Abnorm Psychol, 112*(2), 179-192. doi:10.1037/0021-843X.112.2.179

Buss, K. A., Davidson, R. J., Kalin, N. H., & Goldsmith, H. H. (2004). Context-specific freezing and associated physiological reactivity as a dysregulated fear response. *Developmental Psychology, 40*, 583-594. doi:10.1037/0012-1649.40.4.583

Dadds, M. R., Spence, S. H., Holland, D. E., Barrett, P. M., & Laurens, K. R. (1997). Prevention and early intervention for anxiety disorders: A controlled trial. *J Consult Clin Psychol*, *65*(4), 627-635.

De Wit, C. (1987). *Korte Depressievragenlijst voor Kinderen [Short depression questionnaire for children]*. Amersfoort/Leuven: Acco.

Field, A. P. (2009). *Discovering statistics using SPSS: and sex and drugs and rock 'n' roll (3rd edition)*. London: Sage.

Fox, N. A., Henderson, H. A., Rubin, K. H., Calkins, S. D., & Schmidt, L. A. (2001). Continuity and discontinuity ol behavioral inhibition and exuberance: Psychophysiological and behavioral influences across the first four years of life. *Child Development*, *72*, 1-21. doi:Doi 10.1111/1467-8624.00262

Ge, X., Conger, R. D., & Elder, G. H., Jr. (2001). The relation between puberty and psychological distress in adolescent boy. *Journal of Research on Adolescence, 11*, 49-70.

Goldsmith, H. H., & Rothbart, M. K. (1991). Contemporary instruments for assessing early temperament by questionnaire and in the laboratory. In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament: International perspectives on theory and measurement* (pp. 249-272). New York: Plenum Press.

Hu, X. Z., Lipsky, R. H., Zhu, G., Akhtar, L. A., Taubman, J., Greenberg, B. D., . . . Goldman, D. (2006). Serotonin transporter promoter gain-of-function genotypes are linked to obsessive-compulsive disorder. *American Journal of Human Genetics, 78*(5), 815-826. doi:10.1086/503850

Inderbitzen-Nolan, H. M., & Walters, K. S. (2000). Social Anxiety Scale for Adolescents: Normative data and further evidence of construct validity. *J Clin Child Psychol, 29*(3), 360-371. doi:10.1207/s15374424jccp2903_7

Jansen, W., van de Looij-Jansen, P. M., de Wilde, E. J., & Brug, J. (2008). Feeling fat rather than being fat may be associated with psychological well-being in young dutch adolescents. *J Adolesc Health, 42*(2), 128-136. doi:10.1016/j.jadohealth.2007.07.015

Kendall, P. C. (1994). Treating anxiety disorders in children: Results of a randomized clinical trial. *J Consult Clin Psychol, 62*(1), 100-110.

Kim, K. J., Conger, R. D., Elder, G. H., Jr., & Lorenz, F. O. (2003). Reciprocal influences between stressful life events and adolescent internalizing and externalizing problems. *Child Dev, 74*(1), 127-143.

Kovacs, M. (1981). Rating scales to assess depression in school-aged children. *Acta Paedopsychiatr, 46*(5-6), 305-315.

La Greca, A. M. (1998). *Manual of the Social Anxiety Scales for Children and Adolescents*. Miami, FL: Autor.

La Greca, A. M., & Lopez, N. (1998). Social anxiety among adolescents: Linkages with peer relations and friendships. *J Abnorm Child Psychol, 26*(2), 83-94.

Levy, F., Hawes, D. J., & Johns, A. (2014). Externalizing and Internalizing Comorbidity. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford Handbook of Externalizing Spectrum Disorders*: Oxford Handbooks Online.

Meijer, P., Mellenbergh, G., & de Wit, C. (1986). De Korte Depressie Vragenlijst voor Kinderen (KDVK). *Nederlands Tijdschrift voor de Psychologie, 41*, 364-368.

R Core Team. (2016). *R: A language and environment for statistical computing* R. F. f. S. Computing (Ed.) Retrieved from http://www.r-project.org/

Reynolds, C. R., & Richmond, B. O. (1978). What i think and feel: A revised measure of children's manifest anxiety. *Journal of Abnormal Child Psychology, 6*(2), 271-280.

Reynolds, C. R., & Richmond, B. O. (1985). *Revised Children's Manifest Anxiety Scale. RCMAS Manual*. Los Angeles: Western Psychological Services.

Rylander-Rudqvist, T., Hakansson, N., Tybring, G., & Wolk, A. (2006). Quality and quantity of saliva DNA obtained from the self-administrated oragene method--a pilot study on the cohort of Swedish men. *Cancer Epidemiol Biomarkers Prev, 15*(9), 1742-1745. doi:10.1158/1055-9965.epi-05-0706

Saylor, C. F., Finch, A. J., Jr., Spirito, A., & Bennett, B. (1984). The children's depression inventory: A systematic evaluation of psychometric properties. *Journal of Consulting and Clinical Psychology, 52*(6), 955-967.

Silverman, W. K., Kurtines, W. M., Ginsburg, G. S., Weems, C. F., Lumpkin, P. W., & Carmichael, D. H. (1999). Treating anxiety disorders in children with group cognitive-behaviorial therapy: A randomized clinical trial. *J Consult Clin Psychol, 67*(6), 995-1003.

Singmann, H., Bolker, B., & Westfall, J. (2015). afex: Analysis of factorial experiments. R package version 0.15-2.

Smeekens, S., Riksen-Walraven, J. M., & van Bakel, H. J. A. (2007). Multiple determinants of externalizing behavior in 5-year-olds: A longitudinal model. *Journal of Abnormal Child Psychology, 35*, 347-361. doi:10.1007/s10802-006-9095-y

Smucker, M. R., Craighead, W. E., Craighead, L. W., & Green, B. J. (1986). Normative and reliability data for the Children's Depression Inventory. *J Abnorm Child Psychol, 14*(1), 25-39.

Timbremont, B., & Braet, C. (2002). *Handleiding Children's Depression Inventory*. Lisse: Swets Test Publishers.

van Bakel, H. J. A., & Riksen-Walraven, J. M. (2004). Stress reactivity in 15-month-old infants: Links with infant temperament, cognitive competence, and attachment security. *Developmental Psychobiology, 44*, 157-167. doi:10.1002/dev.20001 van de Looij-Jansen, P. M., Jansen, W., de Wilde, E. J., Donker, M. C. H., & Verhulst, F. C. (2011). Discrepancies between parent-child reports of internalizing problems among preadolescent children: Relationships with gender, ethnic background, and future internalizing problems. *Journal of Early Adolescence, 31*(3), 443-462. doi:10.1177/0272431610366243