CHILD DEVELOPMENT

Child Development, January/February 2014, Volume 85, Number 1, Pages 220–236

Adolescents' Increasing Stress Response to Social Evaluation: Pubertal Effects on Cortisol and Alpha-Amylase During Public Speaking

Esther van den Bos, Mark de Rooij, Anne C. Miers, Caroline L. Bokhorst, and P. Michiel Westenberg
Leiden University

Stress responses to social evaluation are thought to increase during adolescence, which may be due to pubertal maturation. However, empirical evidence is scarce. This study is the first to investigate the relation between pubertal development and biological responses to a social-evaluative stressor longitudinally. Participants performed the Leiden Public Speaking Task twice, with a 2-year interval ($N = 217$; age at Time 1: 8–17 years). The results support an increase in sensitivity to social evaluation during adolescence. The overall cortisol and alpha-amylase responses increased—both between and within participants—and were more strongly related to self-reported pubertal development than to age. The cortisol response shifted from speech delivery toward anticipation. The alpha-amylase response increased in both phases.

It has often been noted that the stress response to social evaluation increases during adolescence. This is suggested by the rising incidence of psychopathologies associated with the stress of social evaluation in this period (Nelson, Leibenluft, McClure, & Pine, 2004), such as depression (Andersen & Teicher, 2008; Nelson et al., 2004; Sisk & Zehr, 2005; Walker, Sabuwalla, & Huot, 2004) and anxiety disorders (Nelson et al., 2004; Walker et al., 2004). Heightened sensitivity to social evaluation is also suggested by subclinical mood changes. Adolescents report increasing awkwardness, embarrassment (Walker et al., 2004), self-consciousness, and social anxiety (Forbes & Dahl, 2010; Westenberg, Drewes, Goedhart, Siebelink, & Treffers, 2004). In contrast to other fears, social fears do not show a decrease from childhood to adolescence, resulting in a relative increase of social fears (Weems & Costa, 2005; Westenberg, Gullone, Bokhorst, Heyne, & King, 2007). The biological changes in puberty have been proposed as a potential explanation of the increasing sensitivity to social evaluation during adolescence (see, e.g., Dahl & Gunnar, 2009; Stroud et al., 2009). Specifically, as discussed elsewhere, puberty involves increasing secretion of gonadal hormones, which affect the maturation of brain structures involved in the stress response. The present research is the first to investigate the relation between pubertal development and the biological response to a social-evaluative stressor in a longitudinal study.

The onset of puberty is marked by gonadarche: the biological process of activating the gonads (Blakemore, Burnett, & Dahl, 2010; Spear, 2000; Walker et al., 2004). Gonadarche begins between the ages of 8 and 14 years in females and a year later in males (Blakemore et al., 2010). The process results in the secretion of gonadal hormones: estrogen (notably estradiol) in females and testosterone in males (Blakemore et al., 2010; Spear, 2000; Walker et al., 2004). The levels of gonadal hormones show a strong rise between the ages of 12 and 20 years (Walker et al., 2004).

Gonadal hormones have both activational (i.e., transient, functional) and organizational (i.e., more or less permanent, structural) effects on the brain (Blakemore et al., 2010; Sisk & Zehr, 2005; Walker et al., 2004). Activational effects of gonadal hormones on the human stress system are under debate. Although animal research has demonstrated that the sensitivity of the hypothalamic–pituitary–adrenal (HPA) axis of the stress system is heightened by estrogen and dampened by testosterone (Viau, 2002), these effects are less clear in humans (Gunnar & Vazquez, 2006; Kudielka & Kirschbaum, 2005).

We thank Sindy Sumter, Laura Compier-de Block, Ellen Middag, Manja Koenders, and Masha Takens for their assistance in data collection. We also thank the children and adolescents who participated in this study.

Correspondence concerning this article should be addressed to Esther van den Bos, Developmental Psychology Unit, Leiden University, PO Box 9555, 2300 RB Leiden, The Netherlands. Electronic mail may be sent to bosejvanden@fsw.leidenuniv.nl.

 0009-3920/2014/8501-0015
DOI: 10.1111/cdev.12118

However, organizational effects seem relevant to the human stress response to social evaluation (Sisk & Zehr, 2005; Spear, 2000).

Organizational effects arise because gonadal hormones affect maturational processes in the brain. Evidence has been found for the influence of gonadal hormones on the maturation of the cortex, the hippocampus, and the amygdala. Maturation of the frontal and parietal lobes of the cortex occurs earlier in females than in males (Blakemore et al., 2010; Sisk & Zehr, 2005). Also, gray (and white) matter volume and density are related to pubertal stage (Blakemore et al., 2010). The hippocampus and the amygdala show sexually dimorphic development (Blakemore et al., 2010; Sisk & Zehr, 2005). In addition, activity of the amygdala in response to social threat is related to pubertal stage (Forbes & Dahl, 2010).

Importantly, the prefrontal cortex, the hippocampus, and the amygdala are involved in regulating the stress system. The hippocampus and the prefrontal cortex exert inhibitory control, whereas the amygdala activates the system (Charmandari, Kino, Souvatzoglou, & Chrousos, 2003; Gunnar & Quevedo, 2007). Maturational changes in these corticolimbic structures, under the influence of gonadal hormones, may result in increasing sensitivity—and hence responding—of the stress system to social evaluation. Two patterns of development have been proposed in the literature. First, sensitivity may show an increase to adult levels (Gunnar & Quevedo, 2007) as these structures reach their adult form. Second, sensitivity may peak during adolescence, showing an increase followed by a decrease (Spear, 2000). Such a peak may result from temporary imbalances in the system due to differences in maturation rate between structures. For example, it has been suggested that social anxiety peaks during adolescence as a result of hyperactivity of the amygdala because of a temporary lack of inhibition by the more slowly maturing prefrontal cortex (Andersen & Teicher, 2008).

Two major components of the stress system can be distinguished. One is the HPA axis. Activation of the HPA axis results in the release of cortisol (in humans). A stress-related increase in cortisol concentrations has effects on the body as well as the brain. In the body, cortisol helps to free up energy for fight or flight (Gunnar & Quevedo, 2007). In the brain, cortisol modulates various aspects of the initial stress response by permitting, stimulating, or suppressing them, to prevent negative long-term consequences (Sapolsky, Romero, & Munck, 2000). Stress levels of cortisol may also prepare the organism for the next stressor (Sapolsky et al., 2000), for example, by sensitizing the amygdala to threat signals (Gunnar & Vazquez, 2006). Finally, cortisol inhibits the further release of cortisol through several negative feedback mechanisms, as a prolonged elevation of cortisol can have negative consequences of its own (Charmandari et al., 2003; Gunnar & Quevedo, 2007).

The other major component of the stress system is the fast-functioning sympathetic adrenomedullary (SAM) system, which is part of the sympathetic branch of the autonomic nervous system (Gunnar & Quevedo, 2007). The SAM system is regulated by the locus coeruleus (LC). The LC releases norepinephrine, which affects vigilance, arousal, and attention (Benarroch, 1993; Gunnar & Quevedo, 2007). The LC also activates structures lower in the brain stem responsible for the sympathetic stress response (Benarroch, 1993; Gunnar & Quevedo, 2007). Sympathetic neurons descending from the brain stem through the spinal cord activate various muscles and glands, including the adrenal glands. In response, the adrenal medulla secretes catecholamines, which are involved in the release of energy for fight or flight (Gunnar & Quevedo, 2007). Alpha-amylase, which is secreted by the salivary glands, is often used as a marker of sympathetic activity (Ehlert, Erni, Hebisch, & Nater, 2006; Granger et al., 2006; Van Stegeren, Rohleder, Everaerd, & Wolf, 2006).

With regard to the stress response to social evaluation, similar patterns of development may be expected for the HPA axis and the SAM system. Although there are some differences in their corticolimbic connections (Gunnar & Quevedo, 2007), both components of the stress system receive input from the prefrontal cortex, the hippocampus, and the amygdala. Moreover, there are reciprocal excitatory connections between the paraventricular nucleus of the hypothalamus and the LC (Charmandari et al., 2003). Because of the effects of gonadal hormones on maturation of the corticolimbic structures that regulate the stress system, the responses to social evaluation by both the HPA axis and the SAM system can be expected to increase during adolescence.

Changes in the stress response to social evaluation have mainly been investigated by exposing participants of different ages to various public performance tasks such as the Trier Social Stress Test (TSST; Kirschbaum, Pirke, & Hellhammer, 1993) and the Leiden Public Speaking Task (LPST; Westenberg et al., 2009). Responses of the HPA axis and the SAM system to these tasks can be demonstrated by heightened levels of salivary cortisol and salivary alpha-amylase (sAA) in an early posttask sample relative to a pretask sample and a late posttask (recovery) sample.

Only a few cross-sectional studies have investigated the effects of age and puberty on the cortisol and alpha-amylase responses to a public performance task between late childhood and adolescence. Although some of the findings are inconsistent, the prediction that the response to social evaluation would increase during adolescence is supported by the overall pattern of results. Gunnar, Wewerka, Frenn, Long, and Griggs (2009) compared the cortisol response to the TSST for 9-, 11-, 13-, and 15-year-olds. They observed a larger cortisol response in 15-year-olds than in 11-year-olds. Moreover, there was a marginally significant correlation between the cortisol response and pubertal stage as measured by self-report on the Pubertal Development Scale (PDS; Petersen, Crockett, Richards, & Boxer, 1988).

Stroud et al. (2009) used the relation between age and Tanner stage, assessed for a subsample of participants either by a physician or by self-report using pictures depicting the Tanner stages, to divide the total sample into children (7–12 years) and adolescents (13–17 years). Participants were randomly assigned to either a peer rejection task, in which the participant had to interact with two confederates of the experimenter and was excluded, or a performance task adapted from the TSST. The peer rejection task yielded a stronger sAA response in adolescents than in children, whereas the performance task yielded a stronger cortisol response in adolescents than in children.

Finally, a study with children and adolescents aged 8–17 years and using the LPST distinguished the cortisol and sAA responses in anticipation of the speech task from the responses to delivering the speech itself (Sumter, Bokhorst, Miers, Van Pelt, & Westenberg, 2010). This was possible because, in this protocol, participants were instructed to prepare a speech on a certain topic a week beforehand. The PDS (Petersen et al., 1988) was used to assess pubertal stage. The results showed an effect of age on the cortisol response to speech anticipation, which was larger in 15- to 17-year-olds than in 9- to 10-year-olds and
11- to 12-year-olds, and marginally larger than in 13- to 14-year-olds. The cortisol response to speech anticipation was also larger at the advanced or postpubertal stage than at the prepubertal stage and marginally larger than at the beginning and midpubertal stages. In contrast, for sAA a significant effect of age was observed on the speech delivery response: It was larger in 13- to 14-year-olds than in 9- to 10-year-olds. The latter effect could not be explained on the basis of pubertal development.

In conclusion, both the cortisol response and the alpha-amylase response to social evaluation seem to increase during adolescence, although developmental effects appeared contingent upon the specific stressor and task phase. Moreover, the role of pubertal development was not directly explored in each study. More evidence is needed to draw firm conclusions about the effects of puberty on the responses of the HPA axis and the SAM system. In addition, all findings were obtained with cross-sectional designs. A longitudinal study is needed to investigate whether or not the sensitivity of the HPA axis and the SAM system increases within individuals during adolescence and puberty and, if so, whether there is an increase to adult levels or an increase followed by a decrease.

It is at present unclear whether the effects of pubertal development on the stress response to social evaluation are similar for both genders. In adults, males sometimes show a larger cortisol response than females (see Kudielka & Kirschbaum, 2005, for a review). Ordaz and Luna (2012) therefore suggested that the increase in the stress response with pubertal development is larger in males than in females. Consistent with this suggestion, Klimes-Dougan, Hastings, Granger, Usher, and Zahn-Waxler (2001) found that interacting with a shy stranger and subsequently delivering an impromptu public speech produced a significant cortisol response in 14- to 16-year-old boys, but not in girls or 11- to 13-year-old boys. However, other studies have reported a larger cortisol response in girls than in boys in late childhood (De Veld, Riksen-Walraven, & De Weerth, 2012) and in midadolescence (Gunnar, Wewerka, et al., 2009). The potential role of pubertal development in the emergence of gender differences in the stress response has not been systematically investigated. For example, a stronger stress response in adolescent girls may be due to their advanced pubertal status.

This study aims to clarify the effects of puberty (and gender) on the normal development of sensitivity to social evaluation during adolescence. The effects of pubertal development on the responses of the HPA axis and the SAM system to a public speaking task were investigated within individuals. To our knowledge, this is the first longitudinal study on the development of sensitivity to social evaluation. Specifically, a cross-sequential design was used. The participants in the cross-sectional study by Sumter et al. (2010), aged 8 to 17 years old, were invited to return to the lab 2 years later and perform the LPST (Westenberg et al., 2009) for a second time. The LPST was specifically designed for a longitudinal study.

Because the task does not take participants by surprise, the procedure can be repeated. In addition, it allows for disentangling the response in anticipation of social evaluation from the response to the actual social-evaluative situation. Salivary cortisol was used as a marker of HPA-axis activity and sAA was used as a marker of SAM-system activity.

For both the overall cortisol response and the overall alpha-amylase response, gender differences were explored first. Subsequently, the developmental hypotheses were tested. On the basis of the effects of gonadal hormones on the maturation of corticolimbic structures regulating the stress system as well as previous findings, it was hypothesized that both the cortisol response and the alpha-amylase response to the LPST would increase with age and pubertal development. The effects of pubertal development were tested both between participants (cross-sectional effects) and within participants (longitudinal effects). Moreover, it was hypothesized that the increase in the cortisol and alpha-amylase responses would be more strongly related to pubertal development than to age. In addition, the possibility of an increase followed by a decrease (i.e., a peak during adolescence) was explored as well as the possibility that the effects of pubertal development on the cortisol and alpha-amylase responses would be moderated by gender. Finally, the effects of gender, age, and pubertal development on the cortisol and alpha-amylase responses were investigated for the two phases of the LPST: speech anticipation and speech delivery. For cortisol, the speech anticipation response in particular was expected to rise with age and pubertal development.

Method

Participants

The data used in this study are part of the Social Anxiety and Normal Development study (SAND; e.g., Miers, Blöte, de Rooij, Bokhorst, & Westenberg, 2013; Westenberg et al., 2009), which was approved by the Leiden University Medical Ethical Committee and carried out in accordance with the Declaration of Helsinki. Parents provided active consent; written assent was obtained from participants themselves.

The SAND study aimed for a normative sample. Participants were recruited through two elementary schools and one secondary school in Leiden, a middle-sized city in the Netherlands. The majority of participants were of Dutch origin: 93.4% of the participants and 87.4% of their parents were born in the Netherlands. The sample included 126 elementary schoolchildren and 169 adolescents from all educational streams in the Dutch school system representing varied levels of intelligence in the whole sample and within all age groups. Present education and living situation of the participants as well as the highest completed education of their parents are summarized in Table 1. Children and adolescents with severe psychological problems or physical illness were excluded from participation. If such problems had been registered at school, students were not invited to participate. To identify any individuals with conditions unknown to the school, participants completed a health and medication history questionnaire probing for treatment by a mental health professional as well as any physical complaints. One participant with a growth disorder, who was treated with daily injections of growth hormone, was excluded from the analyses.

In the SAND study, data were collected in four waves. The LPST was administered in Wave 1 and Wave 3. For clarity, these data collection points will be referred to as Time 1 and Time 2 in this study. At Time 1, there were 295 participants: 151 males (51.2%) and 144 females (48.8%). Their ages ranged from 8 to 17 years. The mean age at Time 1 was 13.10 years (SD = 2.23) for male participants and 13.16 years (SD = 2.32) for female participants.

Table 1
Social Background Characteristics of Participants and Their Parents Assessed at Time 1

Category	Percentage
Living situation participants (n = 289)	
With both biological parents	83.0
With biological mother only	5.9
With biological mother and stepfather	5.2
With each biological parent in alternation	2.8
Other	3.1
Present education participants (n = 295)	
Elementary school	42.7
Prevocational secondary education	6.1
Senior general/preuniversity secondary education (year 1)	11.9
Senior general secondary education (years 2–5)	19.0
Preuniversity education (years 2–6)	20.3
Highest completed education parents (n = 451)	
Elementary school or less	1.11
Prevocational education	17.07
Middle to higher secondary education	9.09
Middle vocational education	18.18
Higher vocational education	26.39
University	26.61
Other	1.55

At Time 2, 2 years later, 217 participants returned: 112 males (51.6%) and 105 females (48.4%). The mean age at Time 2 was 14.9 years (SD = 2.20) for male participants and 15.0 years (SD = 2.28) for female participants. The attrition rate was 26.4%. Participants dropped out of the study for various reasons, including being too busy with school work or hobbies, not enjoying participation at Time 1, and moving out of the area. There was no difference in the distribution of gender, $\chi^2(1) = .060$, $p = .807$, or type of education, $\chi^2(4) = 3.432$, $p = .488$, between participants who returned at Time 2 and participants who did not return. Likewise, there was no difference in mean age, $t(293) < 1$; mean score on the PDS, $t(282) < 1$; task effect (area under the curve [AUC]: increase) on the biological measures: cortisol, $t(232) = -1.045$, $p = .297$; sAA, $t(101,8) < 1$; or mean score on the Social Anxiety Scale, $t(290) = -1.053$. This is consistent with other data from the SAND study indicating that attrition over all four waves of data collection was related to neither social anxiety nor predictors of social anxiety (Miers et al., 2013).

Procedure

The Leiden Public Speaking Task. The LPST was modeled on a classroom presentation, with which participants of all ages have experience. Participants are familiarized with the lab and receive instructions to prepare a speech on a specified topic a week in advance. Participants deliver their speech in front of a projection screen displaying a life-size audience of age peers and a female teacher, who behave neutrally. They are informed that their performance will be recorded and evaluated by peers at a later date. This situation of ambiguous rather than negative social evaluation may be particularly suitable to reveal developmental differences in sensitivity to social evaluation. As participants have no direct control over the way in which their performance will be evaluated, the LPST combines the two characteristics of laboratory procedures that most consistently trigger a response by the HPA axis: social-evaluative threat and uncontrollability (Dickerson & Kemeney, 2004). Previous research has demonstrated a mean cortisol response to the LPST of 2.28 nmol·l^{-1} in 12- to 15-year-olds (Westenberg et al., 2009). This represents a 44% increase over resting levels, which is comparable with the cortisol response observed in other studies with an adolescent sample (see Gunnar, Talge, & Herrera, 2009, for a review). Because participants are fully informed about the upcoming task, the LPST allows for distinguishing between the effects of speech anticipation and speech delivery.

At both Time 1 and Time 2, participants were invited to the lab twice: once for a presession and once for a public speaking session 1 week later. In the presession, they were familiarized with the lab and received instructions to prepare a speech on movies they liked or disliked, in the same way as they would for a presentation at school. In addition, they were instructed to refrain from exercising, smoking, eating, and drinking caffeinated beverages, dairy products, and alcohol 1 hr before the start of the public speaking session. The actual public speaking task consisted of seven parts. First, participants watched a nature video while seated (20 min) and while standing (5 min). Then, participants received instructions, reminding them of the social-evaluative nature of the task (3 min) and they were allowed to rehearse their presentation (5 min). Subsequently the videotape was started and participants watched the audience enter (1 min), after which they delivered their speech (5 min). Finally, there was a posttask recovery period (30 min) during which participants completed assessments and watched another 10 min of the nature video. All sessions started at 2:15 p.m. to minimize diurnal effects. Full details of the task are provided by Westenberg et al. (2009).

Measures

Neuroendocrine measures. Seven saliva samples were used to assess cortisol (nmol·l^{-1}) and alpha-amylase (U·l^{-1}). Saliva samples were collected by passively drooling into plastic vials (IBL-SaliCap®, Hamburg, Germany) directly or through a straw. Sample 1 was taken after the nature video (i.e., pretask sample). After the speech, six samples were taken to account for the fact that individuals differ in the timing of the cortisol response to a stressful event (Gunnar & Talge, 2007). Sample 2 was collected directly after the speech and Sample 3 was taken 10 min later. Samples 4–7 were collected at intervals of 5 min, so that the seventh and last saliva sample was taken at the end of the recovery period (i.e., recovery sample).

The determination of cortisol in saliva was performed with a competitive electrochemiluminescence immunoassay ECLIA using a Modular Analytics E170 immunoassay analyzer from Roche Diagnostics (Mannheim, Germany). The sample volume was at least 20 μl. Cortisol concentrations were only determined for samples from participants who were assessed at both Time 1 and Time 2. Samples from one participant at one time of measurement were

batched together for analysis. All batches were run in a single lot.

Missing values due to insufficient volume ranged between 0.5% and 13.4% of the samples (M = 4.6%). Outliers were conservatively defined as cortisol concentrations of more than 6 *SD* above the mean. As these samples were likely to be contaminated by blood, outliers were removed. Four different participants had outliers in Sample 1 or Sample 7 (179 nmol·l^{-1} in Time 1 Sample 7, 76 nmol·l^{-1} in Time 2 Sample 1, 107 nmol·l^{-1} in Time 2 Sample 1, 102 nmol·l^{-1} in Time 2 Sample 7). Removing these outliers implied that for the time of data collection concerned, the participant could not be included in the statistical analyses. Three other participants each had one outlier in Samples 2–6 (96 nmol·l^{-1} in Time 1 Sample 2, 79 nmol·l^{-1} in Time 2 Sample 3, 85 nmol·l^{-1} in Time 2 Sample 5). These outliers were treated as a missing value for the specific sample. Missing values in Samples 2–6 of the LPST were interpolated by averaging the previous and the next sample. After interpolation, the percentage of missing values ranged from 0 to 6.9 over all samples (M = 3.3%). Thirty-five participants had missing data on one or more samples. In the statistical analyses, the natural logarithm of the cortisol concentrations was used because the data were highly skewed.

The determination of sAA was performed with an enzymatic colorimetric assay using the PNP-G7 substrate at 37°C on a P-module clinical chemistry analyzer (Roche, Mannheim, Germany) in 50-fold diluted saliva samples. Samples from one participant at one time of measurement were batched together. Missing values due to insufficient volume ranged between 0 and 12.2% of the samples (M = 4.6%). Outliers of more than 6 *SD* above the mean were removed. One participant had outliers in six of the seven samples at Time 1 (M = 2,581,475 U·l^{-1}) and another participant had outliers in four of the seven samples at Time 1 (M = 2,609,550), including Sample 7. The Time 1 data from these participants were not included in the statistical analyses. One other participant had an outlier in Time 1 Sample 6 (1,994,050 U·l^{-1}). This outlier was treated as a missing value for the specific sample. Missing values in Samples 2–6 of the LPST were interpolated by averaging the previous and the next sample. After interpolation, the percentage of missing values ranged from 0 to 7.8 over all samples (M = 3.3%). Twenty-seven participants had missing values on one or more samples. In the statistical analyses, the square root of the alpha-amylase concentrations was used because the data were skewed.

Treatment of factors potentially influencing the neuroendocrine measures. At the beginning of the public speaking session, participants filled out a questionnaire on factors potentially influencing the neuroendocrine measures, including current medication usage, eating less than 1 hr before the start of the session, and current phase of the menstrual cycle in girls. Long-term use of medication was assessed with a health and medication history questionnaire, which participants had filled out at home. Sixty-three participants reported using medication for a chronic condition on the health and medication history questionnaire, taking incidentally used medication on the day of the public speaking task or both. The most frequently used types of medication were respiratory medication (n = 15), painkillers (n = 13), and hay fever medication (n = 12). Because the frequencies were too small to control for each kind of medication separately, the use of *any* medication was statistically controlled for. Recent food intake was also statistically controlled for: Nineteen participants had eaten less than 1 hr before the start of the public speaking session.

The aim of controlling for phase of the menstrual cycle is to control for fluctuations in estradiol, which affect the cortisol response. However, the phase of the menstrual cycle can only be determined for girls with a regular cycle, whereas the fluctuations in estradiol begin years before menarche (Shirtcliff, Dahl, & Pollak, 2009). As the phase of the menstrual cycle could not be determined for the majority of girls in this study (73% at Time 1 and 53% at Time 2), it did not seem useful as a control variable.

Pubertal development. Assessment of pubertal status through physical examination by a trained clinician is generally considered the gold standard. In this study, this was not possible because of the large number of assessments required. Therefore, participants filled out a widely used self-report questionnaire: the PDS (Petersen et al., 1988). Although agreement is modest for Tanner staging based on physical examination and Tanner staging based on self-report, moderate to substantial correlations have been observed when the PDS is used as a continuous scale (e.g., Brooks-Gunn, Warren, Rosso, & Gargiulo, 1987; Chan et al., 2010; Shirtcliff et al., 2009). In addition, the use of a self-report measure minimizes the risk of selective drop-out because it is more acceptable to participants than a physical examination (Brooks-Gunn et al., 1987; Petersen et al., 1988; Shirtcliff et al., 2009).

Three items of the PDS were used to compute a pubertal development score. For girls, the items

concerned menarche, pubic hair development, and breast development. For boys, the items concerned voice change, pubic hair development, and facial hair development. These items are the same ones on which pubertal status categories are based according to the PDS manual (Crockett, 1988) because they are more reliable than the two items concerning skin change and growth spurt (Petersen et al., 1988). In this study, too, including these three items led to the most reliable scale. Cronbach's alphas were .86 for boys and .85 for girls at Time 1 and .89 for boys and .82 for girls at Time 2. The individual items were scored on a scale from 1 to 4, except for the item concerning menarche. This item was scored as 1 if the girl had not experienced menarche yet and as 4 if she had (Petersen et al., 1988). The overall pubertal development score was calculated by averaging the ratings on the three items. At Time 1, the PDS was completed by 288 participants (97.6%) and the mean score was 2.1 (SD = 1.0) for boys and 2.7 (SD = 1.1) for girls. At Time 2, the PDS was completed by 216 participants (99.5%) and the mean score was 2.7 (SD = 1.0) for boys and 3.3 (SD = 0.9) for girls. Ten participants had a lower PDS score at Time 2 than at Time 1. As this suggests that at least one of their scores was unreliable (Petersen et al., 1988), these participants were excluded from the analyses.

Analytic strategy

Preliminary analyses. A repeated measures analysis of variance (ANOVA) with sample (1, 2, 3, 4, 5, 6, and 7) and time (1 and 2) as within-subjects factors was performed to establish that the task led to changes in the concentrations of cortisol and sAA in the total sample. All effects were tested in a multivariate approach, which does not assume sphericity. Due to missing values, the sample size was 181 for cortisol and 189 for sAA.

Main analyses. The effects of age and pubertal development were investigated using regression analysis with clustered bootstrap (Cameron, Gelbach, & Miller, 2008; De Rooij, 2012; Harden, 2011) because this technique is suitable for time-varying predictors (De Rooij, 2012). Regression weights were estimated as in standard regression analysis, but the standard errors were derived by bootstrapping. From the total data set, 10,000 samples of the same size as the total set were drawn randomly with replacement. To deal with the dependency between measurements of the same individual, the bootstrap was clustered: Individuals rather than cases were sampled, so that if the individual was assessed at both times, both measurements were included in the sample (De Rooij, 2012). The analyses were run in R 2.5.1 (R Development Core Team, 2007).

Dependent variables. Developmental effects were investigated on three dependent variables for each neuroendocrine measure. First, the overall stress response as measured by the area under the curve with respect to increase (AUCi; Pruessner, Kirschbaum, Meinlschmid, & Hellhammer, 2003) was investigated. AUCi represents the increase in concentration relative to a baseline. Because participants knew beforehand that they had to give a speech, the pretask sample was not a valid baseline. The recovery sample was used as the baseline for cortisol. For sAA, Sample 4 was used as the baseline because concentrations in the recovery sample were similar to those in Samples 4, 5, and 6, suggesting that the fluctuations in the later samples were random. *Decreases* in concentration relative to the baseline (cortisol: 12.5%, sAA: 17.8%) were coded as increases of zero (Pruessner et al., 2003). Second, as in Sumter et al. (2010), the stress responses to two phases of the task were investigated separately: (a) the *speech anticipation response* was defined as the concentration in the pretask sample minus the concentration in the baseline sample and (b) the *speech delivery response* was defined as the maximum concentration in the posttask samples minus the concentration in the pretask sample. Again, decreases were coded as increases of zero.

Models. For each dependent variable, the effects of gender, age, and pubertal development were investigated by testing four regression models. All models included the control variables, recent food intake and medication. Model 1 concerned the main effect of gender (*male* = 1, *female* = 0). Model 2 concerned the effect of age (centralized at the age of the youngest participant at Time 1). Model 3 concerned the hypothesis that the concentrations of the neuroendocrine measures would increase with pubertal development. To test both between- and within-participants effects of pubertal development, the PDS score was split up into an initial level and a change component (De Rooij, 2012; Hedeker & Gibbons, 2006). The PDS score at Time 1 (T1PDS) is a measure of the between-participants effect. The difference between the current PDS score and the PDS score at Time 1 (ΔPDS) is a measure of the within-participants effect. Model 4 concerned the effects of pubertal development while controlling for age.

Two additional models were tested for the overall stress response. Model 5 involved a cross-sectional investigation of the possibility that there would be an increase followed by a decrease in the

concentrations of the neuroendocrine measures with pubertal development. This is an interaction hypothesis, which predicts that concentrations would increase with ΔPDS for participants with lower T1PDS scores and decrease with ΔPDS for participants with higher T1PDS scores. The predictors were gender, T1PDS, ΔPDS, and the interaction between T1PDS and ΔPDS (T1 × ΔPDS). Model 6 concerned the possibility that the effects of pubertal development would be different for males and females. The predictors were gender, T1PDS, ΔPDS, the interactions between gender and T1PDS (G × T1PDS) and gender and ΔPDS (G × ΔPDS). For each model, an index of the goodness of fit was reported: the quasi-likelihood under the independence model criterion (QIC; Pan, 2001). This is a repeated measures variant of Akaike's information criterion. Lower values indicate a better fit.

Results

Descriptives

The mean concentrations of cortisol and alpha-amylase in each of the seven saliva samples collected during the public speaking session are provided in online Supporting Information Table S1. The mean of the overall stress response, the mean response to speech anticipation, and the mean response to speech delivery for cortisol and alpha-amylase are shown in Table 2. Figure 1 illustrates the increase in cortisol concentrations with pubertal development. Figure 2 illustrates the increase in sAA concentrations with pubertal development.

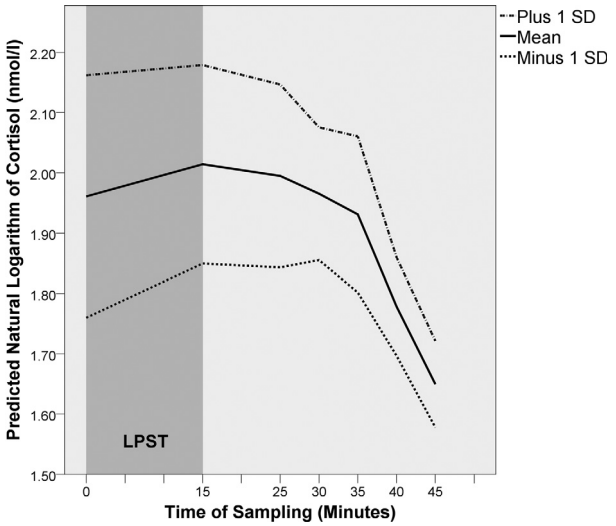

Figure 1. Predicted cortisol concentration in the seven saliva samples taken directly before and 15, 25, 30, 35, 40, and 45 min after the beginning of the Leiden Public Speaking Task (LPST) for scores on the Pubertal Development Scale (PDS; Petersen et al., 1988) 1 *SD* below the mean, at the mean, and 1 *SD* above the mean. PDS score was the only predictor in the regression model (gender differences were not significant). The intercepts and regression weights have been estimated for each of the seven samples, based on the data from both Time 1 and Time 2.

Table 2
Means and Standard Deviations of the Dependent Variables, Response Rates and Recovery Rates for Cortisol and Salivary Alpha-Amylase

	Time 1		Time 2	
Measure	Male	Female	Male	Female
Cortisol				
AUCi	11.7 (10.8)	15.2 (13.0)	13.9 (10.7)	14.9 (11.5)
Speech anticipation	0.32 (0.32)	0.46 (0.45)	0.31 (0.31)	0.41 (0.36)
Speech delivery	0.31 (0.33)	0.23 (0.29)	0.39 (0.44)	0.22 (0.32)
Response rate	95.8%	95.4%	98.9%	95.4%
Recovery rate	68.8%	83.1%	76.3%	81.9%
Alpha-amylase				
AUCi	1,845 (2,089)	2,557 (2,194)	1,823 (1,746)	2,611 (1,936)
Speech anticipation	61 (84)	76 (88)	64 (90)	81 (84)
Speech delivery	89 (111)	99 (117)	88 (92)	98 (108)
Response rate	85.9%	92.6%	87.9%	96.8%
Recovery rate	67.1%	84.1%	71.3%	79.3%

Note. The dependent measures are based on the natural logarithm of the cortisol concentrations ($nmol \cdot l^{-1}$) and the square root of the alpha-amylase concentrations ($U \cdot l^{-1}$) in the saliva samples. Standard deviations are in parentheses. AUCi = area under the curve with respect to increase (Pruessner et al., 2003); Speech anticipation = pretask concentration (directly before start task) minus baseline concentration (cortisol: 45 min after start task; alpha-amylase: 30 min after start task); Speech delivery = posttask peak concentration (cortisol: 25–45 min after start task; alpha-amylase: 25–30 min after start task) minus pretask concentration; Response rate = percentage of participant for whom the baseline concentration was lower than the posttask peak concentration; Recovery rate = percentage of participants for whom the baseline concentration was lower than the pretask concentration.

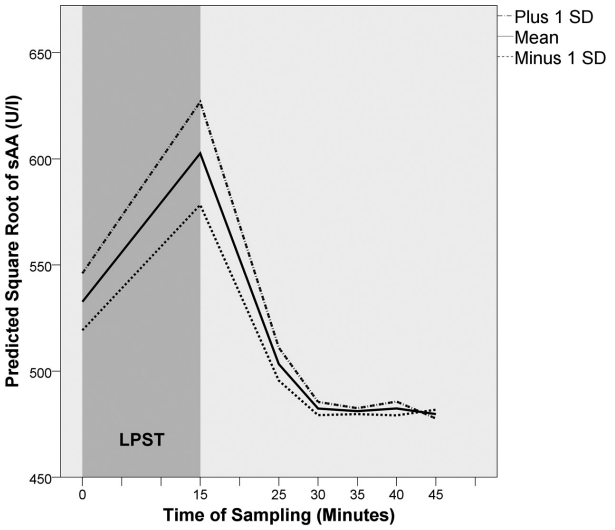

Figure 2. Predicted alpha-amylase concentration in the seven saliva samples taken directly before and 15, 25, 30, 35, 40, and 45 min after the beginning of the Leiden Public Speaking Task (LPST) for scores on the Pubertal Development Scale (PDS; Petersen et al., 1988) 1 *SD* below the mean, at the mean, and 1 *SD* above the mean. PDS score was the only predictor in the regression model (gender differences were not significant). The intercepts and regression weights have been estimated for each of the seven samples, based on the data from both Time 1 and Time 2.

Preliminary Analyses

Cortisol. A repeated measures ANOVA with sample (1, 2, 3, 4, 5, 6, and 7) and time (1 and 2) as within-subjects factors showed a significant main effect of sample, Wilks $\Lambda = .384$, $F(6, 175) = 46.822$, $p < .001$, $\eta_p^2 = .616$, indicating an overall effect of the LPST. Specifically, cortisol concentrations in Sample 1 were significantly *higher* than in Sample 7 ($p < .001$), indicating a speech anticipation effect, whereas cortisol concentrations in Sample 1 were significantly *lower* than in Sample 2 ($p = .004$), indicating a speech delivery effect. In addition, pairwise comparisons with Bonferroni correction showed that cortisol concentrations were higher in the posttask samples than at recovery: cortisol concentrations in Samples 2, 3, 4, 5, and 6 were significantly higher than in Sample 7 ($p < .001$ for all comparisons).

The repeated measures ANOVA also showed a significant main effect of time, Wilks $\Lambda = .827$, $F(1, 180) = 37.699$, $p < .001$, $\eta_p^2 = .173$; cortisol concentrations averaged over all samples increased from Time 1 ($M = 1.75$, $SD = .52$) to Time 2 ($M = 1.98$, $SD = .53$). The interaction between sample and time was marginally significant, Wilks $\Lambda = .940$, $F(6, 175) = 1.865$, $p = .089$, $\eta_p^2 = .060$. Although cortisol concentrations increased from Time 1 to Time 2 in all samples, the increase was particularly large for later posttask samples. Planned comparisons (least significant difference test) following a repeated measures ANOVA on the difference scores of each sample showed that the increase was larger for Sample 3 ($p = .032$), Sample 5 ($p = .004$), and Sample 6 ($p = .031$) than for Sample 2. For Sample 4, the increase was marginally larger than for Sample 2 ($p = .067$). For Sample 5, the increase was also larger than for Samples 1 ($p = .029$) and 7 ($p = .032$). These results indicate that the duration of the stress response increased from Time 1 to Time 2.

Alpha-amylase. A repeated measures ANOVA with sample (1, 2, 3, 4, 5, 6, and 7) and time (1 and 2) as within-subjects factors showed a significant main effect of sample, Wilks $\Lambda = .444$, $F(6, 183) = 38.243$, $p < .001$, $\eta_p^2 = .556$. Specifically, sAA concentrations in Sample 1 were significantly *higher* than in Sample 7 ($p < .001$), indicating a speech anticipation effect, whereas sAA concentrations in Sample 1 were significantly *lower* than in Sample 2 ($p < .001$), indicating a speech delivery effect. Pairwise comparisons with Bonferroni correction showed that sAA concentrations were higher shortly after the LPST (Sample 2) than at recovery (Sample 7, $p < .001$). There were no significant differences between Samples 4, 5, 6, and 7.

The repeated measures ANOVA also showed a significant main effect of time, Wilks $\Lambda = .975$, $F(1, 188) = 4.873$, $p = .028$, $\eta_p^2 = .025$; sAA concentrations averaged over all samples decreased from Time 1 ($M = 503.96$, $SD = 186.73$) to Time 2 ($M = 488.14$, $SD = 174.80$). There was no significant interaction between sample and time, Wilks $\Lambda = .950$, $F(6, 183) = 1.613$, $p = .146$.

Overall Stress Responses: Gender and Developmental Effects

Cortisol. Table 3 shows the QIC for each model and the regression weights with 95% confidence intervals for the theoretically relevant predictors. Model 1 indicated a significant effect of gender. This can be concluded from the 95% confidence interval [−5.0 to −0.01], which does not include zero. Because *male* was coded as 1 and *female* was coded as 0, the negative regression weight indicates that AUCi was larger for girls than for boys. Model 2 showed significant effects of age and gender: AUCi increased with age and was larger for girls than for boys. Model 3 showed that AUCi increased with pubertal development. Both T1PDS and ΔPDS were significant predictors of AUCi, indicating that the effect of pubertal development was significant both

Table 3
Goodness of Fit (QIC) and Regression Weights of the Predictors With 95% Confidence Intervals for Six Regression Models Predicting the Overall Cortisol Response as Measured by the Area Under the Curve With Respect to Increase

M	QIC	Gender	Age	T1PDS	ΔPDS	T1 × ΔPDS	G × T1PDS	G × ΔPDS
1	2835	−2.5 [−5.0 to −0.01]*	—	—	—	—	—	—
2	2815	−2.4 [−4.7 to −0.03]*	1.2 [0.7 to 1.6]*	—	—	—	—	—
3	2812	−0.5 [−2.9 to 1.9]	—	3.0 [1.9 to 4.2]*	2.2 [0.1 to 4.3]*	—	—	—
4	2814	−0.8 [−3.5 to 1.9]	0.3 [−0.6 to 1.2]	2.5 [0.3 to 4.6]*	1.8 [−0.9 to 4.4]	—	—	—
5	2814	−0.4 [−2.8 to 2.0]	—	2.9 [1.7 to 4.0]*	1.6 [−1.3 to 4.5]	1.1 [−4.2 to 2.1]	—	—
6	2816	−0.5 [−2.9 to 2.0]	—	3.0 [1.9 to 4.2]*	2.3 [0.1 to 4.4]*	—	0.4 [−2.7 to 1.9]	−0.2 [−4.5 to 4.0]

Note. The analyses were performed on the natural logarithm of the cortisol concentrations (nmol·l^{-1}). Eating less than 1 hr before the public speaking session and the use of medication were controlled for. Gender was coded 1 = *male*, 0 = *female*. Age was centralized at the age of the youngest participant at Time 1. M = model; QIC = quasi-likelihood under the independence model information criterion; T1PDS = score on the Pubertal Development Scale (PDS; Petersen et al., 1988) at Time 1; ΔPDS = difference score on the PDS (current—Time 1); T1 × ΔPDS = interaction between T1PDS and ΔPDS; G × T1PDS = interaction between gender and T1PDS; G × ΔPDS = interaction between gender and ΔPDS.
*$p < .05$.

between participants and within participants. The main effect of gender was no longer significant when pubertal development was controlled for. In Model 4, including age and pubertal development, only T1PDS remained a significant (positive) predictor of AUCi. ΔPDS, gender, and age were not significant. This implies that the overall cortisol response is more strongly related to pubertal development than to age.

Concerning the additional hypotheses, Model 5 showed no evidence of an increase followed by a decrease in AUCi. T1PDS was a significant (positive) predictor, but the interaction between T1PDS and ΔPDS, and gender were not. Model 6 showed neither any interactions between gender and pubertal development nor a main effect of gender when pubertal development was controlled for, but T1PDS and ΔPDS were significant (positive) predictors of AUCi.

Recent food intake was a significant negative predictor of the AUCi for cortisol in Models 2, 3, 4, and 5. The use of medication was not a significant predictor in any model. The QICs show that Model 3 had the best fit to the data, indicating that AUCi was most efficiently predicted by pubertal development. The overall cortisol response to the LPST was larger as participants had higher PDS scores at Time 1 and larger increases in PDS score from Time 1 to Time 2.

Alpha-amylase. Table 4 shows the QIC for each model and the regression weights with 95% confidence intervals for the theoretically relevant predictors. Model 1 showed a significant effect of gender: AUCi was larger for girls than for boys. Model 2 showed significant effects of age and gender: AUCi increased with age and was larger for girls than for boys. Model 3 showed that AUCi increased with pubertal development. Both T1PDS and ΔPDS were significant predictors of AUCi, indicating that the effect of pubertal development was significant both between participants and within participants. The main effect of gender was no longer significant when pubertal development was controlled for. In Model 4, including age and pubertal development, only ΔPDS remained a significant (positive) predictor of AUCi. T1PDS, gender, and age were not significant. This implies that the overall sAA response is more strongly related to pubertal development than to age.

Concerning the additional hypotheses, Model 5 showed no evidence of an increase followed by a decrease in AUCi. None of the predictors reached significance. Model 6 showed neither any interactions between gender and pubertal development

Table 4
Goodness of Fit (QIC) and Regression Weights of the Predictors With 95% Confidence Intervals for Six Regression Models Predicting the Overall Salivary Alpha-Amylase Response as Measured by the Area Under the Curve With Respect to Increase

M	QIC	Gender	Age	T1PDS	ΔPDS	T1 × ΔPDS	G × T1PDS	G × ΔPDS
1	8197	−506 [−961 to −50]*	—	—	—	—	—	—
2	8194	−493 [−947 to −38]*	99 [13 to 185]*	—	—	—	—	—
3	8192	−341 [−822 to 139]	—	269 [50 to 488]*	484 [63 to 905]*	—	—	—
4	8194	−315 [−830 to 201]	22 [−200 to 157]	313 [−125 to 750]	519 [0.7 to 1,036]*	—	—	—
5	8193	−317 [−802 to 168]	—	227 [−12 to 467]	314 [−200 to 828]	298 [−995 to 399]	—	—
6	8195	−335 [−815 to 144]	—	279 [52 to 505]*	501 [75 to 926]*	—	93 [−550 to 363]	−416 [−1,278 to 446]

Note. The analyses were performed on the square root of the alpha-amylase concentrations ($U \cdot l^{-1}$). Eating less than 1 hr before the public speaking session and the use of medication were controlled for. Gender was coded 1 = *male*, 0 = *female*. Age was centralized at the age of the youngest participant at Time 1. M = model; QIC = quasi-likelihood under the independence model information criterion; T1PDS = score on the Pubertal Development Scale (PDS; Petersen et al., 1988) at Time 1; ΔPDS = difference score on the PDS (current—Time 1); T1 × ΔPDS = interaction between T1PDS and ΔPDS; G × T1PDS = interaction between gender and T1PDS; G × ΔPDS = interaction between gender and ΔPDS.
*$p < .05$.

nor a main effect of gender, but T1PDS and ΔPDS were significant (positive) predictors of AUCi.

Recent food intake and the use of medication did not significantly predict the AUCi for sAA in any of the models. The QICs show that Model 3 had the best fit to the data, indicating that AUCi was most efficiently predicted by pubertal development. The overall sAA response to the LPST was larger as participants had higher PDS scores at Time 1 and larger increases in PDS score from Time 1 to Time 2.

Responses to Speech Anticipation: Gender and Developmental Effects

Cortisol. Table 5 shows the QIC for each model and the regression weights with 95% confidence intervals for the theoretically relevant predictors. Model 1 showed a significant effect of gender: The speech anticipation response was larger for girls than for boys. Model 2 showed significant effects of age and gender: The speech anticipation response increased with age and was larger for girls than for boys. Model 3 showed that the speech anticipation response increased with pubertal development. T1PDS was a significant (positive) predictor, but ΔPDS was not, indicating that the effect of pubertal development was significant between participants, but not within participants. The main effect of gender was no longer significant when pubertal development was included. In Model 4, including age and pubertal development, only T1PDS was a significant (positive) predictor of the speech anticipation response, whereas ΔPDS, gender, and age were not significant. This suggests that the speech anticipation response is more strongly related to pubertal development than to age.

Recent food intake was a significant negative predictor of the speech anticipation effect for cortisol in all models. The use of medication was not a significant predictor in any model. The QICs show that Model 3 had the best fit to the data, indicating that the speech anticipation response was most efficiently predicted by pubertal development. The cortisol response to speech anticipation was larger when participants had higher PDS scores at Time 1.

Alpha-amylase. Table 5 shows the QIC for each model and the regression weights with 95% confidence intervals for the theoretically relevant predictors. Models 1 and 2 indicated that the speech anticipation response could not be predicted from gender alone or from age and gender. Model 3 indicated that the speech anticipation response increased with pubertal development. T1PDS was significant, but ΔPDS was not, indicating that the effect of pubertal development was significant

Table 5
Goodness of Fit (QIC) and Regression Weights of the Predictors With 95% Confidence Intervals for Four Regression Models Predicting the Speech Anticipation Effect for Cortisol and Salivary Alpha-Amylase

M	QIC	Gender	Age	T1PDS	ΔPDS
Cortisol					
1	317	−0.11 [−0.20 to −0.03]*	—	—	—
2	298	−0.11 [−0.19 to −0.03]*	0.04 [0.02 to 0.05]*	—	—
3	285	−0.04 [−0.12 to 0.04]	—	0.11 [0.07 to 0.14]*	0.01 [−0.05 to 0.07]
4	287	−0.03 [−0.12 to 0.06]	−0.01 [−0.04 to 0.02]	0.12 [0.05 to 0.19]*	0.02 [−0.06 to 0.10]
Alpha-amylase					
1	5369	−3.1 [−22.2 to 16.0]	—	—	—
2	5367	−2.7 [−21.7 to 16.2]	3.5 [−.04 to 7.5]	—	—
3	5367	3.5 [−15.9 to 22.9]	—	10.3 [1.2 to 19.5]*	12.3 [−4.8 to 29.4]
4	5369	4.9 [−16.4 to 26.3]	−1.2 [−8.7 to 6.3]	12.7 [−4.4 to 29.8]	14.2 [−7.1 to 35.4]

Note. The analyses were performed on the natural logarithm of the cortisol concentrations ($nmol \cdot l^{-1}$) and on the square root of the alpha-amylase concentrations ($U \cdot l^{-1}$). Eating less than 1 hr before the public speaking session and the use of medication were controlled for. Gender was coded 1 = *male*, 0 = *female*. Age was centralized at the age of the youngest participant at Time 1. M = model; QIC = quasi-likelihood under the independence model information criterion; T1PDS = score on the Pubertal Development Scale (PDS; Petersen et al., 1988) at Time 1; ΔPDS = difference score on the PDS (current—Time 1).
*$p < .05$.

between participants, but not within participants. The main effect of gender when pubertal development was controlled for was not significant either. In Model 4, including age and pubertal development, none of the predictors remained significant. This suggests that age and pubertal development largely explained the same variance. In addition, the QICs show that there was no difference in goodness of fit between Models 2 and 3. However, only Model 3 contained a significant predictor: The sAA response to speech anticipation was larger when participants had higher PDS scores at Time 1. Recent food intake and the use of medication did not significantly predict the speech anticipation effect for alpha-amylase in any of the models.

Responses to Speech Delivery

Cortisol. Table 6 shows the QIC for each model and the regression weights with 95% confidence intervals for the theoretically relevant predictors. Model 1 showed a significant effect of gender: The speech delivery response was larger for boys than for girls. Model 2 did not show a significant effect of age, but the effect of gender remained significant. Model 3 indicated that the speech delivery response

Table 6
Goodness of Fit (QIC) and Regression Weights of the Predictors With 95% Confidence Intervals for Four Regression Models Predicting the Speech Delivery Effect for Cortisol and Salivary Alpha-Amylase

M	QIC	Gender	Age	T1PDS	ΔPDS
Cortisol					
1	308	0.12 [0.04 to 0.19]*	—	—	—
2	307	0.11 [0.04 to 0.19]*	−0.01 [−0.03 to 0.002]	—	—
3	303	0.08 [0.002 to 0.16]*	—	−0.05 [−0.08 to −0.01]*	0.04 [−0.04 to 0.11]
4	305	0.08 [−0.003 to 0.15]	0.01 [−0.03 to 0.04]	−0.06 [−0.13 to 0.01]	0.03 [−0.07 to 0.13]
Alpha-amylase					
1	5575	−9.8 [−32.5 to 13.0]	—	—	—
2	5571	−9.2 [−31.7 to 13.4]	4.9 [1.0 to 8.9]*	—	—
3	5573	−2.7 [−24.7 to 19.3]	—	11.2 [1.2 to 21.1]*	14.5 [−4.4 to 33.4]
4	5575	−5.9 [−29.3 to 17.5]	2.7 [−5.6 to 10.9]	5.7 [−14.7 to 26.2]	10.3 [−12.2 to 32.8]

Note. The analyses were performed on the natural logarithm of the cortisol concentrations ($nmol \cdot l^{-1}$) and on the square root of the alpha-amylase concentrations ($U \cdot l^{-1}$). Eating less than 1 hr before the public speaking session and the use of medication were controlled for. Gender was coded 1 = *male*, 0 = *female*. Age was centralized at the age of the youngest participant at Time 1. M = model; QIC = quasi-likelihood under the independence model information criterion; T1PDS = score on the Pubertal Development Scale (PDS; Petersen et al., 1988) at Time 1; ΔPDS = difference score on the PDS (current—Time 1).
*$p < .05$.

decreased with pubertal development. T1PDS was a significant predictor, but ΔPDS was not, indicating that the effect of pubertal development was significant between participants but not within participants. The main effect of gender remained significant when pubertal development was included. In Model 4, including age and pubertal development, none of the predictors remained significant. This suggests that age and pubertal development largely explained the same variance. However, the finding that the speech delivery response could not be predicted from age in Model 2, whereas it could be predicted from pubertal development in Model 3 implies that it is more strongly related to pubertal development than to age. The QICs show that Model 3 had the best fit to the data. The cortisol response to speech delivery was smaller when participants had higher PDS scores at Time 1. Recent food intake and the use of medication did not significantly predict the speech delivery effect for cortisol in any of the models.

Alpha-amylase. Table 6 shows the QIC for each model and the regression weights with 95% confidence intervals for the theoretically relevant predictors. Model 1 showed no main effect of gender. Model 2 indicated that the speech delivery response increased with age. The effect of gender was not significant. Model 3 showed that the speech delivery response increased with pubertal development. T1PDS was significant, but ΔPDS was not, indicating that the effect of pubertal development was significant between participants but not within participants. The main effect of gender was also not significant. In Model 4, including age and pubertal development, none of the predictors remained significant. This suggests that, to a large extent, age and pubertal development explained the same variance.

The QICs show that Model 2 had the best fit to the data, indicating that the speech delivery response was most efficiently predicted by age. The sAA response to speech delivery was larger when participants were older. However, the effects of age and pubertal development could not be disentangled. The use of medication was a significant negative predictor of the speech delivery response for alpha-amylase in all models. Recent food intake was not a significant predictor in any model.

Discussion

Overall Stress Responses

This study is the first to demonstrate a consistent effect of self-reported pubertal development on two key components of the human stress system across a 2-year period. The overall response of the HPA axis (cortisol) and the SAM system (sAA) to a public speaking task increased with self-reported pubertal development. For both systems this effect was observed between individuals (i.e., cross-sectional effect of level of pubertal development) and within individuals (i.e., longitudinal effect of pubertal change). Indeed, for both systems the effect of self-reported pubertal development was stronger than the effect of age. For alpha-amylase, the average concentration across all samples decreased over time, but this decrease probably reflected a change in resting levels. The alpha-amylase response increased with age and pubertal development.

This pattern of findings supports the hypothesis that pubertal development increases adolescents' sensitivity to social evaluation (Gunnar & Quevedo, 2007; Spear, 2000). The results appear consistent with a gradual increase in the stress response to adult levels as the corticolimbic structures regulating the stress system grow into their adult forms under the influence of gonadal hormones (Gunnar & Quevedo, 2007). For neither the HPA axis nor the SAM system was there any evidence that the effect of pubertal development between Time 1 and Time 2 was moderated by the initial level of pubertal development. Hence, the present findings do not support a puberty-related peak in the stress response due to temporary imbalances in the input from different corticolimbic structures maturing at different rates (Andersen & Teicher, 2008). This conclusion is supported by the findings of cross-sectional studies showing that the stress response to social evaluation does not differ between early to middle adolescence and young adulthood (Kudielka, Buske-Kirschbaum, Hellhammer, & Kirschbaum, 2004), whereas the stress response differs significantly between late childhood and young adulthood (Strahler, Mueller, Rosenloecher, Kirschbaum, & Rohleder, 2010; Yim, Granger, & Quas, 2010).

There was no evidence for consistent gender differences in this study. When the effect of gender was considered in isolation, the overall cortisol and alpha-amylase responses were larger for girls than for boys. This is consistent with studies showing a larger cortisol response in girls than in boys in late childhood and midadolescence (De Veld et al., 2012; Gunnar, Wewerka, et al., 2009). However, in this study, the effects of gender disappeared when pubertal development was controlled for, suggesting that girls showed a larger stress response because of their higher level of pubertal development than boys in the same age range.

Although gender differences have not been demonstrated consistently, some studies with adult participants have reported a larger cortisol response to social evaluation in males than in females (for overviews, see Dickerson & Kemeney, 2004; Kudielka & Kirschbaum, 2005). Ordaz and Luna (2012) therefore hypothesized that males would show a stronger cortisol response increase with pubertal development than females. Despite its substantial sample and approximately equal male-to-female ratio, this study did not find support for this hypothesis. Although gender-specific effects of pubertal development on the cortisol response may be revealed with real-life stressors, the results of this study suggest that gender effects are limited, at least in adolescence.

Responses to Speech Anticipation and Speech Delivery

Separate investigation of the stress response to two phases of the task—speech anticipation and speech delivery—resulted in basically the same pattern of results for cortisol and sAA during speech anticipation: a puberty-related increase. For speech delivery, the findings were inconsistent: a puberty-related increase for sAA and a decrease for cortisol. Furthermore, for both task phases the effect of pubertal development was stronger than the age effect, but it proved difficult to disentangle both effects, and the effect of pubertal development was obtained between but not within individuals. This relative lack of clarity may be because difference scores are less sensitive than the AUC (Pruessner et al., 2003). The AUC is sensitive to both height and duration of the stress response and might be more suitable for detecting developmental change within individuals.

At the same time, following Sumter et al. (2010), distinguishing between the two phases proved to be particularly relevant for the response of the HPA axis. The contradictory effect of pubertal development on cortisol—a positive relation for speech anticipation and a negative relation for speech delivery—suggests that pubertal development strengthens the response to speech anticipation and the latter replaces the response to speech delivery. As can be seen in Figure 1, more mature participants do not only reach higher peak levels but they also approach these levels already before the actual public speaking task. Moreover, preliminary analyses showed that the *duration* of the cortisol response increased from Time 1 to Time 2. Overall these findings suggest that in the course of pubertal development the stress response builds up at an earlier moment and is maintained during the delivery of the speech.

Increasing sensitivity in anticipation of social evaluation during adolescence was also observed with a different paradigm in a study on neurodevelopmental changes in brain activity (Gunther Moor, Van Leijenhorst, Rombouts, Crone, & Van der Molen, 2010). In this study, children, adolescents, and adults were presented with faces of age peers and they had to predict whether or not each person would like them. Activation of the ventral medial prefrontal cortex and the striatum increased with age when people were *waiting* for the verdict and not when the verdict was received. Unfortunately, the effect of pubertal development was not investigated.

Conclusions

In conclusion, this study indicates that the HPA axis and the SAM system become increasingly sensitive to social evaluation during adolescence. The relatively strong relation with pubertal development implies that a certain increase in sensitivity to social evaluation during adolescence is inevitable. This may reflect an adaptive mechanism as it makes adolescents more susceptible to social evaluation at a moment in life when they have to approach and adjust to people from outside their own family environment (Spear, 2000). At the same time, heightened stress sensitivity may add to vulnerability to psychopathology. Thus, the results of this study offer a possible explanation of the rise in mood, anxiety, and psychotic disorders in adolescence (Andersen & Teicher, 2008; Nelson et al., 2004; Sisk & Zehr, 2005; Walker et al., 2004). In addition, these data on the normal development of sensitivity to social evaluation may be useful for identifying adolescents who deviate from the normal pattern and are at risk for developing a disorder.

A limitation of this study is that we used a self-report measure of pubertal development rather than an assessment by a trained physician. Although this study found positive relations between self-reported pubertal development and the responses of the HPA axis and the SAM system to a social-evaluative situation, stronger results may be obtained using physician ratings of pubertal development, as these are more reliable (e.g., Brooks-Gunn et al., 1987; Shirtcliff et al., 2009).

Moreover, the pubertal development score was based on both physical changes associated with gonadal hormones (facial hair, voice change, menarche, breast development) and physical changes associated with adrenal hormones (pubic hair). Thus, this study does not conclusively demonstrate

a positive relation between levels of gonadal hormones and physiological responding to social evaluation. In addition, more research is needed to test the hypothesis that such a relation would be mediated by organizational effects on the corticolimbic structures regulating the stress system. Although gonadal hormones are known to affect the maturation of brain structures such as the cortex, the hippocampus, and the amygdala in humans, the evidence is mainly indirect (Blakemore et al., 2010; Sisk & Zehr, 2005). Likewise, the relation between the development of corticolimbic structures and the development of the responses of the HPA axis and the SAM system is inferred from the connections between these structures (Spear, 2000; Stroud et al., 2009). To test the hypothesis directly, longitudinal research combining brain imaging with the measurement of gonadal hormones on the one hand and physiological stress responses on the other will be needed.

Another potential limitation of this study is that only sAA was used as a measure of SAM-system activity. First, there is some debate on the meaning of elevated alpha-amylase concentrations. Although alpha-amylase is generally considered to be a marker of SAM-system activity (Ehlert et al., 2006; Granger et al., 2006; Strahler et al., 2010; Van Stegeren et al., 2006), it has also been suggested to reflect cognitive effort (Yim et al., 2010) or arousal associated with negative as well as positive emotions (Adam, Hoyt, & Granger, 2011). Second, it is at present unclear whether or not the same pattern of development of the sympathetic stress response can be seen with different markers of SAM-system activity. The present finding of an age- or puberty-related increase in the alpha-amylase response to social evaluation is consistent with other studies using alpha-amylase as a marker (Strahler et al., 2010; Stroud et al., 2009; Sumter et al., 2010; Yim et al., 2010). However, the results of two studies using cardiovascular markers of sympathetic activity showed no relation with age (Hollenstein, McNeely, Eastabrook, Mackey, & Flynn, 2012) or even suggested a decrease (Gunnar, Wewerka, et al., 2009). Further research comparing different measures of SAM-system activity is needed to establish whether age and pubertal development have a uniform effect on the sympathetic stress response or affect the various physiological functions controlled by the autonomic nervous system in different ways or to different degrees.

Finally, adolescence is characterized by various cognitive, emotional, and social changes that may or may not be related to pubertal development. For example, although the increase in the stress response in anticipation of a social-evaluative situation was more strongly related to pubertal development than to age, it may be even more strongly related to the development of planning abilities. Similarly, although this study found no evidence of a puberty-related peak in the stress responses, adolescents may be more sensitive to social evaluation than children and adults for another reason. Adolescents experience a shift in social orientation from parents to peers (Nelson et al., 2004; Spear, 2000), which may be accompanied by conflicts with parents (Spear, 2000) as well as uncertainty concerning acceptance by peers (Nelson et al., 2004), whereas the social relationships of children and adults may be more stable. Further research is needed to investigate the interplay of physical, cognitive, emotional, and social factors on the development of sensitivity to social evaluation.

References

Adam, E. K., Hoyt, L. T., & Granger, D. A. (2011). Diurnal alpha amylase patterns in adolescents: Associations with puberty and momentary mood states. *Biological Psychology*, *88*, 170–173. doi:10.1016/j.biopsycho.2011.07.007

Andersen, S. L., & Teicher, M. H. (2008). Stress, sensitive periods and maturational events in adolescent depression. *Trends in Neurosciences*, *31*, 183–191. doi:10.1016/j.tins.2008.01.004

Benarroch, E. E. (1993). The central autonomic network: Functional organization, dysfunction, and perspective. *Mayo Clinical Proceedings*, *68*, 988–1001.

Blakemore, S. J., Burnett, S., & Dahl, R. E. (2010). The role of puberty in the developing adolescent brain. *Human Brain Mapping*, *31*, 926–933. doi:10.1002/hbm.21052

Brooks-Gunn, J., Warren, M. P., Rosso, J., & Gargiulo, J. (1987). Validity of self-report measures of girls' pubertal status. *Child Development*, *58*, 829–841. doi:10.1111/j.1467-8624.1987.tb01423.x

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap based improvements for inference with clustered errors. *Review of Economics and Statistics*, *90*, 414–427. doi:10.1162/rest.90.3.414

Chan, N. T. P., Sung, R. Y. T., Nelson, E. A. S., So, H. K., Tse, Y. K., & Kong, A. P. S. (2010). Measurement of pubertal status with a Chinese self-report pubertal development scale. *Maternal and Child Health Journal*, *14*, 466–473. doi:10.1007/s10995-009-0481-2

Charmandari, E., Kino, T., Souvatzoglou, E., & Chrousos, G. P. (2003). Pediatric stress: Hormonal mediators and human development. *Hormone Research*, *59*, 161–179. doi:10.1159/000069325

Crockett, L. J. (1988). *Pubertal Development Scale: Pubertal categories*. Unpublished manuscript.

Dahl, R. E., & Gunnar, M. R. (2009). Heightened stress responsiveness and emotional reactivity during pubertal maturation: Implications for psychopathology. *Development and Psychopathology*, *21*, 1–6. doi:10.1017/S0954579409000017

De Rooij, M. (2012). *Standard regression models for repeated measurement data*. Manuscript submitted for publication.

De Veld, D. M. J., Riksen-Walraven, M., & De Weerth, C. (2012). The relation between emotion regulation strategies and physiological stress responses in middle childhood. *Psychoneuroendocrinology*, *37*, 1309–1319. doi: 10.1016/j.psyneuen.2012.01.004

Dickerson, S. S., & Kemeney, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*, 355–391. doi:10.1037/0033-2909.130.3.355

Ehlert, U., Erni, K., Hebisch, G., & Nater, U. (2006). Salivary α-amylase levels after yohimbine challenge in healthy men. *Journal of Clinical Endocrinology & Metabolism*, *91*, 5130–5133. doi:10.1210/jc.2006-0461

Forbes, E. E., & Dahl, R. E. (2010). Pubertal development and behavior: Hormonal activation of social and motivational tendencies. *Brain and Cognition*, *72*, 66–72. doi:10.1016/j.bandc.2009.10.007

Granger, D. A., Kivlighan, K. T., Blair, C., El Sheikh, M., Mize, J., Lisonbee, J. A., et al. (2006). Integrating the measurement of salivary α-amylase into studies of child health, development and social relationships. *Journal of Social and Personal Relationships*, *23*, 267–290. doi: 10.1177/0265407506062479

Gunnar, M., & Quevedo, K. (2007). The neurobiology of stress and development. *Annual Review of Psychology*, *58*, 145–173. doi:10.1146/annurev.psych.58.110405.085605

Gunnar, M. R., & Talge, N. M. (2007). Neuroendocrine measures in developmental research. In L. A. Schmidt & S. J. Segalowitz (Eds.), *Developmental psychophysiology* (pp. 343–366). Cambridge, UK: Cambridge University Press.

Gunnar, M. R., Talge, N. M., & Herrera, A. (2009). Stressor paradigms in developmental studies: What does and does not work to produce mean increases in salivary cortisol. *Psychoneuroendocrinology*, *34*, 953–967. doi:10.1016/j.psyneuen.2009.02.010

Gunnar, M. R., & Vazquez, D. (2006). Stress neurobiology and developmental psychopathology. In D. Cichetti & D. Cohen (Eds.), *Developmental psychopathology: Developmental neuroscience* (Vol. 2, 2nd ed., pp. 533–577). New York: Wiley.

Gunnar, M. R., Wewerka, S., Frenn, K., Long, J. D., & Griggs, C. (2009). Developmental changes in hypothalamus-pituitary-adrenal activity over the transition to adolescence: Normative changes and associations with puberty. *Development and Psychopathology*, *21*, 69–85. doi:10.1017/S0954579409000054

Gunther Moor, B., Van Leijenhorst, L., Rombouts, S. A. R. B., Crone, E. A., & Van der Molen, M. W. (2010). Do you like me? Neural correlates of social evaluation and developmental trajectories. *Social Neuroscience*, *5*, 461–482.

Harden, J. (2011). A bootstrap method for conducting statistical inference with clustered data. *State Politics & Policy Quarterly*, *11*, 223–246. doi:10.1177/1532440011406233

Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis*. New York: Wiley.

Hollenstein, T., McNeely, A., Eastabrook, J., Mackey, A., & Flynn, J. (2012). Sympathetic and parasympathetic responses to social stress across adolescence. *Developmental Psychobiology*, *54*, 207–214. doi: 10.1002/dev.20582

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The "Trier Social Stress Test": A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*, 76–81. doi:10.1159/000119004

Klimes-Dougan, B., Hastings, P. D., Granger, D. A., Usher, B. A., & Zahn-Waxler, C. (2001). Adrenocortical activity in at-risk and normally developing adolescents: Individual differences in salivary cortisol basal levels, diurnal variation, and responses to social challenges. *Development and Psychopathology*, *13*, 695–719. doi: 10.1017/S0954579401003157

Kudielka, B. M., Buske-Kirschbaum, A., Hellhammer, D. H., & Kirschbaum, C. (2004). HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: Impact of age and gender. *Psychoneuroendocrinology*, *29*, 83–98. doi: 10.1016/S0306-4530(02)00146-4

Kudielka, B. M., & Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: A review. *Biological Psychology*, *69*, 113–132. doi:10.1016/j.biopsycho.2004.11.009

Miers, A. C., Blöte, A. W., de Rooij, M., Bokhorst, C. L., & Westenberg, P. M. (2013). Trajectories of social anxiety during adolescence and relations with cognition, social competence, and temperament. *Journal of Abnormal Child Psychology*, *41*, 97–110. doi:10.1007/S10802-012-9651-6

Nelson, E. E., Leibenluft, E., McClure, E. B., & Pine, D. S. (2004). The social re-orientation of adolescence: A neuroscience perspective on the process and its relation to psychopathology. *Psychological Medicine*, *35*, 163–174. doi:10.1017/S0033291704003915

Ordaz, S., & Luna, B. (2012). Sex differences in physiological reactivity to acute psychosocial stress in adolescence. *Psychoneuroendocrinology*, *37*, 1135–1157. doi:10.1016/j.psyneuen.2012.01.002

Pan, W. (2001). Akaike's information criterion in generalized estimating equations. *Biometrics*, *57*, 120–125. doi:10.1111/j.0006-341X.2001.00120.x

Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity and initial norms. *Journal of Youth and Adolescence*, *17*, 117–133. doi:10.1007/BF01537962

Pruessner, J. C., Kirschbaum, C., Meinlschmid, G., & Hellhammer, D. H. (2003). Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology*, *28*, 916–931. doi: 10.1016/S0306-4530(02)00108-7

R Development Core Team. (2007). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Sapolsky, R. M., Romero, L. M., & Munck, A. U. (2000). How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews*, *21*, 55–89. doi:10.1210/er.21.1.55

Shirtcliff, E. A., Dahl, R. E., & Pollak, S. D. (2009). Pubertal development: Correspondence between hormonal and physical development. *Child Development*, *80*, 327–337. doi:10.1111/j.1467-8624.2009.01263.x

Sisk, C. L., & Zehr, J. L. (2005). Pubertal hormones organize the adolescent brain and behavior. *Frontiers in Neuroendocrinology*, *26*, 163–174. doi:10.1016/j.yfrne.2005.10.003

Spear, L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neuroscience and Biobehavioral Reviews*, *24*, 417–463. doi:10.1016/S0149-7634(00)00014-2

Strahler, J., Mueller, A., Rosenloecher, F., Kirschbaum, C., & Rohleder, N. (2010). Salivary α-amylase stress reactivity across different age groups. *Psychophysiology*, *47*, 587–595. doi:10.1111/j.1469-8986.2009.00957.x

Stroud, L. R., Foster, E., Papandonatos, G. D., Handwerger, K., Granger, D. A., Kivlighan, K. T., et al. (2009). Stress response and the adolescent transition: Performance versus peer rejection stressors. *Development and Psychopathology*, *21*, 47–68. doi:10.1017/S0954579409000042

Sumter, S. R., Bokhorst, C. L., Miers, A. C., Van Pelt, J., & Westenberg, P. M. (2010). Age and puberty differences in stress responses during a public speaking task: Do adolescents grow more sensitive to social evaluation? *Psychoneuroendocrinology*, *35*, 1510–1516. doi:10.1016/j.psyneuen.2010.05.004

Van Stegeren, A., Rohleder, N., Everaerd, W., & Wolf, O. T. (2006). Salivary alpha amylase as marker for adrenergic activity during stress: Effect of betablockade. *Psychoneuroendocrinology*, *31*, 137–141. doi:10.1016/j.psyneuen.2005.05.012

Viau, V. (2002). Functional cross-talk between the hypothalamic-pituitary-gonadal and adrenal axes. *Journal of Neuroendocrinology*, *14*, 506–513. doi:10.1046/j.1365-2826.2002.00798.x

Walker, E. F., Sabuwalla, Z., & Huot, R. (2004). Pubertal neuromaturation, stress sensitivity, and psychopathology. *Development and Psychopathology*, *16*, 807–824. doi:10.1017/S0954579404040027

Weems, C. F., & Costa, N. (2005). Developmental differences in the expression of childhood anxiety symptoms and fear. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*, 656–663. doi:10.1097/01.chi.0000162583.25829.4b

Westenberg, P. M., Bokhorst, C. L., Miers, A. C., Sumter, S. R., Kallen, V. L., Van Pelt, J., et al. (2009). A prepared speech in front of a pre-recorded audience: Subjective, physiological, and neuroendocrine responses to the Leiden Public Speaking Task. *Biological Psychology*, *82*, 116–124. doi:10.1016/j.biopsycho.2009.06.005

Westenberg, P. M., Drewes, M. J., Goedhart, A. W., Siebelink, B. M., & Treffers, P. D. A. (2004). A developmental analysis of self-reported fears in late childhood through mid-adolescence: Social-evaluative fears on the rise? *Journal of Child Psychology and Psychiatry*, *45*, 481–495.

Westenberg, P. M., Gullone, E., Bokhorst, C. L., Heyne, D. A., & King, N. J. (2007). Social evaluation fear in childhood and adolescence: Normative developmental course and continuity of individual differences. *British Journal of Developmental Psychology*, *25*, 471–483. doi:10.1348/026151006X173099

Yim, I. S., Granger, D. A., & Quas, J. A. (2010). Children's and adults salivary alpha amylase responses to a laboratory stressor and to verbal recall of the stressor. *Developmental Psychobiology*, *52*, 598–602. doi:10.1002/dev.20453

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Table S1. Means and Standard Deviations of Cortisol and Alpha-Amylase Concentrations in Each Saliva Sample at Time 1 and Time 2 for Male and Female Participants.