

Cortisol reactions to a social evaluative paradigm in 5- and 6-year-old children

C. de Weerth, M. A. C. Zijlmans, S. Mack & R. Beijers

To cite this article: C. de Weerth, M. A. C. Zijlmans, S. Mack & R. Beijers (2013) Cortisol reactions to a social evaluative paradigm in 5- and 6-year-old children, *Stress*, 16:1, 65-72, DOI: [10.3109/10253890.2012.684112](https://doi.org/10.3109/10253890.2012.684112)

To link to this article: <https://doi.org/10.3109/10253890.2012.684112>



Published online: 23 May 2012.



Submit your article to this journal [↗](#)



Article views: 617



View related articles [↗](#)



Citing articles: 3 View citing articles [↗](#)

Cortisol reactions to a social evaluative paradigm in 5- and 6-year-old children

C. DE WEERTH, M. A. C. ZIJLMANS, S. MACK, & R. BEIJERS

Department of Developmental Psychology, Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

(Received 2 November 2011; revised 5 April 2012; accepted 6 April 2012)

Abstract

The goal of the present study was to develop a stress paradigm to elicit cortisol secretory responses in a group of 5- and 6-year-old children as a whole. To this end, we tested a paradigm containing elements of social evaluative threat, unpredictability and uncontrollability, and with a duration of 20 min. The Children's Reactions to Evaluation Stress Test is composed of three short tasks that children have to perform in front of a judge. The tasks are rigged so as to provoke (partial) failure in the child's performance. Participants were 42 children ($M = 68.0$ months, $SD = 4.3$). Six saliva samples were taken during the testing session to obtain cortisol measurements of baseline concentrations, stress reactivity, and recovery. Our findings showed that this paradigm was effective in provoking a significant increase in salivary cortisol concentration in the group as a whole, with no effects of possible confounders (child's sex, age or school, parental educational level, time of testing, sex of experimenter, and sex of judge). The mean cortisol concentration increase for the group was 127.5% ($SD = 190.9$); 61% of the children could be classified as reactors (mean increase of 214%, $SD = 201.5$), and 39% as non-reactors (mean decrease of 7.8%, $SD = 16.8$). To our knowledge, this is the first study in this age group that shows a significant cortisol response for the group as a whole to a standardized laboratory paradigm. As such, this paradigm is a promising tool to be used in future research on early life interactions between physiology and psychology.

Keywords: *HPA axis, laboratory stress paradigm, salivary cortisol, social evaluation, stress reactivity, young children*

Introduction

The hypothalamic pituitary adrenal (HPA) axis, with its end product, the glucocorticoid hormone cortisol, is an important regulator of our physiological reactions to stressors. Although humans are regularly faced with stress and need this physiological response to survive, abnormal reactivity of the HPA axis is related to mental disorders (Fairchild et al. 2008; Petrowski et al. 2010) and poorer physical health (Segerstrom and Miller 2004). However, issues of causality in the relationships between abnormal HPA axis reactivity and (psycho)pathology remain unclear. As the HPA axis is sensitive to environmental factors and changes over time (Gunnar and Quevedo 2007), it is important to examine HPA axis reactivity at several time points during development to detect gradual change. Only then we can begin to understand the complex interaction between physiology and

psychology in early life that can lead to mental and physical disorders in adulthood.

The experience of stress is elicited by a physical or psychological challenging situation (Gunnar and Quevedo 2007). Different stimuli or situations are experienced as challenging at different ages. For example, babies react with an increase in cortisol release to physical stressors such as vaccinations (Jansen et al. 2010), whereas children above 7 years of age react to psychological stressors that contain social evaluation, such as a public speaking task (Gunnar et al. 2009a,b). Apparently, different stressors are needed to examine HPA axis reactivity changes over time. Gunnar et al. (2009a) showed in a review that for children between 5 and 6 years of age various paradigms have been used, including frustration tasks (Luby et al. 2003), rigged failure tasks (Lewis and Ramsay 2002), rigged competition (Donzella et al. 2000), strange events (Quas et al. 2004), and

separation tasks (Luby et al. 2003). However, none of these paradigms was able to elicit a significant cortisol increase in the group as a whole.

The goal of this study was to develop a paradigm that would elicit cortisol elevations in 5- and 6-year-olds in the group as a whole. Requirements were that the paradigm would be easy to carry out, and not overly stressful (i.e. not distress the children, and followed by a quick recovery). To this end, we designed a paradigm [the Children's Reactions to Evaluation Stress Test (CREST)] with characteristics of social evaluation, unpredictability, and uncontrollability. These elements have been found to be stressful in children aged 8 years or older and in adults (Dickerson and Kemeny 2004; Gunnar et al. 2009a). Another element that has been suggested to be stressful for young children is failing publicly on a task that they believe even younger children could complete successfully (Gunnar et al. 2009a). This element of forced failure was therefore also included in the paradigm. In short, the CREST was composed of three tasks that children had to perform in front of a judge. The tasks were rigged so as to provoke (partial) failure in the child's performance. Our hypothesis was that this paradigm would produce significant cortisol increases in a group of 5- and 6-year-olds as a whole. The possible effects of the following confounders were explored: child's sex, age and school, parental educational level, time of testing, sex of experimenter, and sex of judge.

Methods

Participants

Children were recruited from four regular primary schools in the Netherlands: one in the city of Nijmegen and three in nearby villages. In the Netherlands, children attend primary school from the age of 4 years onwards, with the first 2 years consisting of kindergarten. Parents of children attending their second year of school were approached by letter ($N = 179$). In this letter, the study was described, and parents were invited to enroll their child if he/she wished to participate. Forty-six applications were received, from which three children were excluded due to a clinically referred diagnosis and/or daily use of medication affecting cortisol secretion. The data of a fourth child were excluded due to physical illness (earache) on the testing day. The final group therefore consisted of 42 children (20 boys and 22 girls), with ages ranging between 57.0 and 75.9 months ($M = 68.0$ months, $SD = 4.3$). Written informed consent was obtained from all parents, and the study was approved by the Ethical Committee of the Faculty of Social Sciences, Radboud University Nijmegen, which follows the Helsinki Declaration.

Procedure

Prior to the testing session, the parents were telephoned and informed about the procedure of the experiment in detail. They were asked to tell their child that the experimenter would ask them to do a couple of tasks during the session, but not to tell their child the exact nature of the tasks. Furthermore, all parents were asked to fill out a short demographics questionnaire including maternal and paternal educational level, and parents' and child's country of origin. The paradigm was carried out in a mobile laboratory (research van) parked next to the child's school. All the children were working in their own classroom prior to the testing session. This had the advantage of creating a relatively standard pre-test situation. Furthermore, because school is a familiar environment where children are accustomed to performing tasks, being taken to the mobile laboratory for the testing session was expected to avoid the possible arousal and anticipation stress of coming to the university laboratory for a testing session.

The testing took place in the afternoon, starting between 13:15 and 15:30 h, in order to avoid the circadian morning peaks of cortisol secretion (Kudielka and Wüst 2010). Furthermore, because food and physical activity can influence cortisol secretion, the teacher was asked not to allow the child to eat or do physical activities 30 min prior to the start of the testing session. Finally, the teacher was also asked to tell the child that the experimenter would ask him/her to do a couple of tasks during the session, but not to reveal the exact nature of the tasks.

Experimental protocol: CREST

The paradigm had a duration of 20 min, and consisted of three tasks (15 min), followed by a period of stress from an anticipated evaluation (5 min). The timeline of the experimental paradigm is included in Figure 2. The children were tested by two researchers: an experimenter and a judge. The experimenter had the role of explaining the tasks and guiding the child through the testing session. If the child showed signs of distress, the experimenter gave extra support to the child during the tasks. The judge had the role of evaluating the child's performance on the tasks. The age of both the experimenter and the judge ranged between 20 and 26 years.

After collecting the child in the classroom, the experimenter brought the child to the research van. To check whether children were really naive about the CREST, the experimenter asked them what the parents, teacher, or other children had told them about the session. There were no indications that any of the children had prior knowledge about the procedure. Subsequently, the experimenter motivated the child by presenting four presents (tissue, used

eraser, bubble blower, and kaleidoscope) and letting the child decide which present he/she liked most and which present he/she liked least. The child was then told that he/she had to perform some tasks in front of a judge. It was emphasized that the judge would decide which present the child deserved (i.e. the most liked or the least liked), depending on his/her task performance. Next, the judge entered the van and took a seat in front of the child and the experimenter.

During the *first task*, the child was asked to stand still in front of the judge, so that the judge could evaluate how still the child was able to stand. The child was told that it was extremely important not to move; otherwise an alarm would go off. This alarm consisted of a timer clicked onto their pocket with two wires that were fastened with velcro strips to their wrist and ankle. First, the experimenter demonstrated the task with a 30-s performance, and the alarm did not go off. Next, the child had to stand still in front of the judge for 60 s. However, irrespective of the child's movements, the alarm went off on two preprogrammed times (after 20 and 40 s). Each time the alarm went off, the child was reminded by the judge that it was very important not to move.

During the *second task*, the child listened to a story about animals (3 min) recorded on an mp3 player. Five seconds of silence followed each animal name, with eight animal names in total. The child was asked to imitate the sound produced by the animal, every time he/she heard the name of an animal. The judge would evaluate the child's performance by showing a green card each time the sound was perfect. Irrespective of the performance, the child was only shown a green card in three out of eight animal sounds (first, second, and sixth sound).

In the *third task*, the child was asked to make a tower of empty soft drink cans identical to the one shown by the experimenter. The experimenter then uncovered an example tower which was invisibly glued, and consisted of a pyramid of cans on their side [four, three, two, and one can(s) in each layer]. The judge told the child that the task was very easy for children to perform, and should therefore work out fine. In reality, when the child tried to build the tower, the cans kept on rolling away making the task impossible. After 3 min, the judge instructed the child that he/she had to stop building the tower.

Hence, all three parts of the test contained elements of *social evaluation*, as the judge was observing and judging the child's behavior throughout. The three tasks were also *uncontrollable*, as the child's behavior was insufficient to control the outcome (i.e. alarm went off twice, five animal sounds were not judged as perfect, and the tower of cans was impossible to build). Finally, *unpredictability* was ensured by the child's prior lack of knowledge of the nature of the tasks, by the unpredictability of the judge's

response, and by the alarm suddenly going off without the child moving.

After completion of the tasks, the child was told that the judge was going to decide if he/she deserved the most liked or the least liked present and the judge went away for 5 min. This was done to elicit stress due to an anticipated evaluation in the child. During these 5 min, the child waited and chose a drawing of a popular cartoon to color in on his/her own. After 5 min, the judge returned and informed the child that he/she had performed very well on the tasks and definitely deserved the nicest present. This was the end of the paradigm. Subsequently, the child was debriefed by telling him/her that the experimenters had been pulling his/her leg and by showing the child how the test was rigged. After the debriefing, the child stayed in the research van with the experimenter for an additional 35 min, during which the child colored a drawing and watched an educational TV program. This time was used to provide a calm and enjoyable experience after the stressfulness of the test, and for collecting the remaining saliva samples. The child received a letter for his/her parents, including a comprehensive debriefing and a short report on the child's reactions to the paradigm. Finally, the child was asked not to tell the content of the test to other children, and was returned to the classroom or picked up by parents if school was already out.

Coercion was not used to initiate or complete testing. One child showed some signs of distress during the paradigm, but these signs disappeared when the experimenter gave the child extra support. No other children showed behavioral signs of distress, and all were able to complete the session.

Cortisol measurements

Six saliva samples were taken to obtain cortisol measurements of baseline, stress reactivity, and recovery concentrations. Two samples measured baseline concentrations: one was taken just before the stress test (C1; pre-stress) and one 15 min after starting the stress test (C2; pre-response; after completion of the three tasks, but before the period of stress due to an anticipated evaluation; previously it has been shown that salivary cortisol increases are not seen at this time point after the start of stressor exposure, Dickerson and Kemeny 2004). To assess stress reactivity, two samples were obtained 25 (C3) and 35 (C4) min after the beginning of the stress test. Lastly, two samples obtained at 45 and 60 min after the beginning of the stress test were used as recovery measurements (C5 and C6, respectively). Eye sponges (BD Visispear™, Waltham, MA) were used as saliva sampling devices (de Weerth et al. 2007). The participant had to put an eye sponge in his/her mouth for approximately 1 min. After that, the eye sponge was transferred into a plastic tube.

These samples were taken to the laboratory and centrifuged for 10 min at 3948g, and then stored in a freezer (-25°C) before analyses. Samples were analyzed by the Psychobiology Laboratory at the University of Trier, Germany. The inter-assay coefficients of variation ranged between 7.1% and 9%; the lower detection limit was 0.173 nmol/L for a 50 μl saliva sample.

Confounders

The following confounders were included because they can possibly influence cortisol reactivity: child's sex, age and school, parental educational level, time of testing, sex of experimenter, and sex of judge. Child sex was not matched to the sex of the experimenter or the judge. Parental educational level was computed by averaging the mother's and father's highest completed educational level (secondary or intermediate vocational education, higher vocational education, and university).

Statistical analyses

Logarithmic transformations were applied to skewed data and data were checked for outliers. All cortisol concentrations were skewed and hence were log transformed. One outlier was detected in the variable

Table I. Descriptive statistics of demographic characteristics and confounders ($N = 42$).

Variables	Value
Maternal educational level (%)	
Secondary or intermediate vocational education	45.2
Higher vocational education	28.6
University	26.2
Paternal educational level (%)	
Secondary or intermediate vocational education	45.2
Higher vocational education	31.0
University	23.8
Child age in months, mean $\pm$ SD (range)	68.0 $\pm$ 4.3 (57.0–75.9)
Child sex (%)	
Boys	47.6
Girls	52.4
Time of testing, hours, mean $\pm$ SD (range)	13:48 h $\pm$ 35 min (13:14–15:28 h)
School (%)	
School 1	26.2
School 2	40.5
School 3	21.4
School 4	11.9
Sex of experimenter (%)	
Male	14.3
Female	85.7
Sex of judge (%)	
Male	71.4
Female	28.6

“percentage increase.” This outlier was replaced by the variable mean plus three times its standard deviation (Hasings et al. 1947).¹

The percentage increase in cortisol was calculated as highest stress reactivity value (C3 or C4) minus lowest baseline value (C1 or C2) divided by lowest baseline value (C1 or C2) and multiplied by 100. To calculate how many children had reacted to the paradigm with relevant cortisol increases, we used the definition of reaction by Schuetze et al. (2008), namely an increase in cortisol concentration from baseline to peak of at least twice the rate of error in the assay and twice the lower limit of assay sensitivity. Hence, children showing an increase from baseline (lowest value of samples C1 and C2) to peak (highest value of samples C3 and C4) of 18% and an actual increase from baseline to peak of 0.35 nmol/L were considered to have reacted to the paradigm with an increase in cortisol release.

To analyze the course of the cortisol concentrations (baseline–stress reactivity–recovery), a repeated measures of analyses of variance (RM-ANOVA) was computed. Greenhouse–Geisser corrections were applied where appropriate. To analyze whether this paradigm induced a significant increase between the baseline cortisol concentration (lowest value of samples C1 and C2) and stress reactivity concentration (highest of samples C3 and C4), a paired samples *t*-test was computed. A paired samples *t*-test was also computed between the stress reactivity concentration (highest of samples C3 and C4) and the recovery concentration (lowest of samples C5 and C6) to analyze whether the children's cortisol concentration significantly decreased after the stress period.

Results

Preliminary analyses

Table I shows the demographic characteristics and potential confounders of the sample.

Correlations between the actual increase in cortisol (highest peak concentration minus lowest baseline concentration) and the possible confounders child's sex and age, parental educational level, time of testing, sex of experimenter, and sex of judge were non-significant (Table II). The actual increase in cortisol concentration also did not differ between the children from different schools ($F_{3,37} = 0.41, p > 0.05$). These confounders were therefore not taken into account in the main analyses.

Main analyses

To analyze the course of the cortisol concentrations during the testing session, a RM-ANOVA was computed. The analysis revealed a significant quadratic

Table II. Correlations between the actual increase in salivary cortisol concentration (nmol/L) and possible confounders.

	Child's sex	Child's age ^a	Parental educational level ^a	Time of testing ^a	Sex of experimenter	Sex of judge
Δ Cortisol	0.04	0.06	-0.12	0.13	0.09	-0.22

^a Pearson correlations; others are Spearman correlations. All non-significant.

time effect for the repeated measurements of cortisol ($F(1.96, 78.5) = 7.54, p < 0.001$). Paired samples t -tests showed a significant difference between baseline and peak cortisol concentrations ($t_{40} = -5.05, p < 0.001, \eta^2 = 0.39$), indicating that the paradigm induced a significant increase in children's salivary cortisol concentrations for the group as a whole (Figure 1). Furthermore, there was a significant difference between peak and recovery cortisol concentrations ($t_{41} = 7.88, p < 0.001, \eta^2 = 0.60$), indicating that the children's cortisol concentrations recovered after the paradigm.

Although the group as a whole showed an increase in salivary cortisol concentration as a reaction to the paradigm, important inter-individual variability in reactivity was observed. The percentage change in cortisol concentration ranged between -52.0% and 748% ($M = 127.5\%$, $SD = 190.9$). Based on the definition of responders of Schuetze et al. (2008), of the total group 61% were found to have reacted significantly ($t = -5.67, p < 0.001, df = 24$), showing a mean increase in cortisol of 214%, whereas 39% were found not to have reacted, showing a non-significant 7.8% mean decrease in salivary cortisol concentration ($t = 1.75, p > 0.05, df = 15$; Figure 2).

Discussion

In this study, we evaluated cortisol reactions to a social evaluative paradigm in 5- and 6-year-old children. Our findings showed that this paradigm was effective in provoking a significant increase in salivary cortisol concentration in the group as a whole. The possible confounders, children's sex, age and school, parental educational level, time of testing, sex of experimenter, and sex of judge, showed no relation with children's cortisol reactivity. This implies that the CREST could be a suitable paradigm to elicit a cortisol response for both 5- and 6-year-old boys and girls from different socioeconomic backgrounds and schools, and that the paradigm can be carried out by both male and female researchers.

The results show that a paradigm including elements of social evaluative threat, uncontrollability and unpredictability, is adequate for inducing a cortisol response in 5- and 6-year-olds. The mean increase for the whole group in salivary cortisol concentrations from baseline to the peak stress response was 127.5%. This mean cortisol increase is comparable to the mean increase obtained by stress tests for children above 8 years of age, including

the Trier Social Stress Test (TSST) for children (Buske-Kirschbaum et al. 1997, ca. 125% mean increase; Kudielka et al. 2004 ca. 120% mean increase; Yim et al. 2010, ca. 150% mean increase), and to the mean increase obtained in a recent study in which a new effective laboratory stressor was developed for 3-year-olds (Kryski et al. 2011; ca. 80% mean increase). It is important to note that, although effective, the CREST was not extremely stressful. The children showed no behavioral signs of being overly stressed or distressed by the paradigm, and displayed a rapid post-stressor cortisol recovery (Figure 1), as well as signs of being relaxed and in a good mood after the debriefing.

To our knowledge, this is the first paradigm showing that social evaluative threat, strengthened with elements of unpredictability and uncontrollability, results in a cortisol response in children aged 5 and 6 years. This is in line with research showing that this combination of elements instead of the separate elements produces cortisol increases in adults and older children (Dickerson and Kemeny 2004; Gunnar et al. 2009a). The strengths of the paradigm are first that a "judge" observed the child's performance during the three tasks. This researcher "judge" was previously unknown to the child. This judging element most probably gave salience to the social

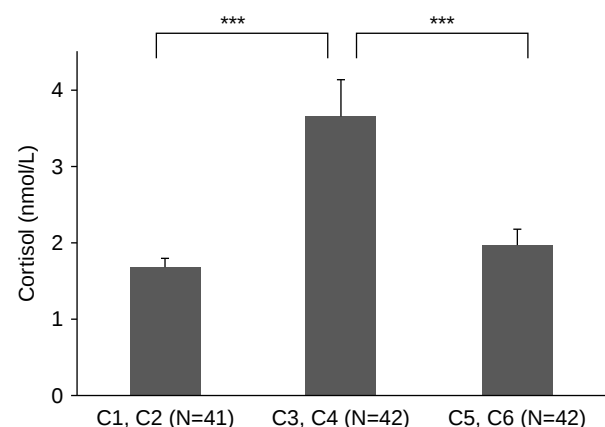


Figure 1. Salivary cortisol concentrations (mean $\pm$ SEM) over the CREST paradigm: baseline concentration (lowest of samples C1 and C2), stress concentration (highest of samples C3 and C4), and recovery concentration (lowest of samples C5 and C6) for the group as a whole. N = number of children/samples for each collection point. Paired samples t -tests showed a significant difference in cortisol concentrations between baseline and stress measurements, and a significant difference between stress and recovery measurements ($***p < 0.001$).

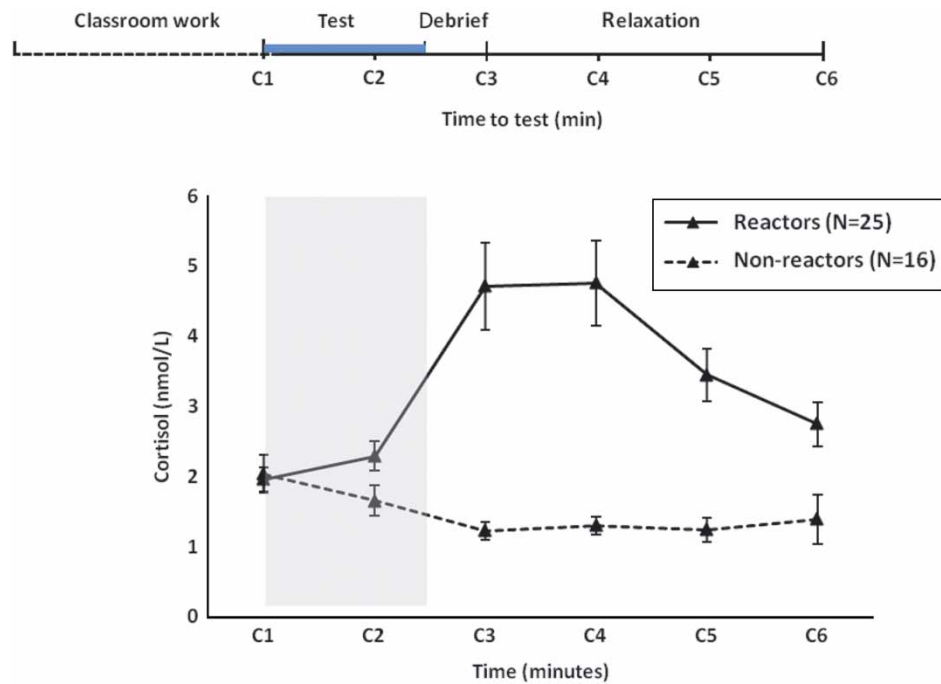


Figure 2. Top: detailed timeline of the CREST paradigm. Saliva sampling points for cortisol assessment are marked as C1, C2 (baseline), C3, C4 (stress reaction), C5, C6 (recovery). The thick bar indicates the stress phase (20 min: 15 min of performance stress plus 5 min of stress due to an anticipated evaluation). Bottom: salivary cortisol concentrations (mean $\pm$ SEM) over the CREST (gray box) for reactors, $N = 25$ children; mean peak increase in salivary cortisol concentration [mean increase from lowest baseline sample (C1/C2) to highest stress sample (C3/C4)] = 214%; $p < 0.001$, paired t -test; and non-reactors, $N = 16$ children; mean decrease in cortisol [mean change from lowest baseline sample (C1/C2) to highest stress sample (C3/C4)] = 7.8%; not significant, paired t -test.

evaluative threat factor of the procedure, producing a high ego involvement. Second, the content of the tasks was unknown beforehand, and the judging and alarm functioning were unrelated to the child's movements, when asked to stand still, or quality of the animal noises imitated, making the testing session *unpredictable*. Third, the forced failure on all three tasks made the procedure a highly *uncontrollable* situation for the children. Moreover, the tasks appeared to be relatively easy beforehand, and in the can tower building task the children were also informed that the tower was very easy for other children to build.

Another, perhaps complementary, explanation for the cortisol response to this social evaluative paradigm could be that it combines several stressful tasks into one paradigm. While some of the earlier stress paradigms were based on a single task (Luby et al. 2003), the CREST combines three different tasks and a period of stress due to an anticipated evaluation into a single paradigm. It is possible that the use of these stressful tasks individually would not lead to significant increases in cortisol release, and that only a combination of tasks leads to a situation that is stressful enough to provoke a cortisol reaction. It is also possible that children differ in which task(s) they experience as (more) stressful. For example, some children could become more stressed by not being able to stand still, while others become more stressed by not being correctly judged for producing perfect

animal noises or building an apparently easy can tower. By combining different tasks, we may have increased the probability that all children were stressed by at least one of them. For example, in a study using the MacArthur Story Stem Battery (Von Klitzing et al. 2003), 5-year-olds had to complete a story about stressful everyday life events with the use of play figures. Although this stressor was chosen because it includes high ego involvement, it only led to a significant cortisol increase in girls (Hatzinger et al. 2007). In another study focusing on rigged failure, most of the children did not show a cortisol response to the task. However, from the 15% of the children who did show a cortisol increase, all but one was male (Donzella et al. 2000). This indicates that a test that combines tasks with both high ego involvement and rigged failure (as well as other stressful elements) could be stressful for both girls and boys. In sum, this study's succession of three short stressful tasks followed by a period of stress due to an anticipated evaluation could therefore at least partly explain the effectiveness of this paradigm for the group as a whole. The tasks could produce additive effects of stress and/or, due to the different nature of the tasks, they could ensure that most of the children were stressed by one of the tasks.

Finally, a further component of the paradigm that may have stimulated the children's cortisol responses could be that by increasing the children's motivation

for participating by promising a preferred gift if they performed well and an unwanted gift if they did not perform well, children were more motivated to succeed in the test. Being motivated to succeed would in turn make the performance on the tasks more relevant and the forced failure therefore more stressful to the children.

A question is why a subgroup of participants (39%) failed to show a cortisol *increase* as a result of the paradigm. The initial cortisol values of this non-reacting group were not significantly different from those of the reacting group, so differences in pre-stress levels were not behind the differences in reaction. In this context, it is important to note that having 39% of participants not reacting to a stress test is not unusual. Although often not reported, most stress tests for children do not produce a cortisol reaction in all individuals (Klimes-Dougan et al. 2001; Gunnar et al. 2009a). Given the nature of our study group, i.e. young children, our paradigm was limited by ethical constraints and therefore designed not to be overly stressful. It is therefore quite possible that some children were simply not stressed by the paradigm. Nonetheless, many other possible explanations exist for the large inter-individual differences in reactivity. Possible candidate factors that could play a role in individual cortisol reactions are temperament, experience with social evaluative situations, and adverse experiences in early development. Also, increases versus decreases in cortisol release as reactions to stressors may represent different biological profiles that are related to individual differences in emotional reactions to stressors. According to Moons et al. (2010), in a study with the TSST in adults, anger reactions to psychosocial stress would be related to increases in cortisol release, while fear reactions would be related to decreases in cortisol. Measuring children's emotions as a reaction to the CREST paradigm, and relating them to their cortisol increases, decreases, or lack of change in cortisol release, is a logical next step in this line of research. Future research will therefore hopefully shed more light on why some children "fail" to react with a cortisol increase to the CREST while others show increases of more than 700%.

Limitations and future directions

This study found that a social evaluative paradigm, strengthened with elements of unpredictability and uncontrollability, elicits a cortisol response in young children. Therefore, the paradigm could be a promising tool for future research on cortisol reactivity and recovery in 5- and 6-year-olds. However, there are some limitations. First, although we asked the teacher not to allow the child to eat or do physical activities 30 min prior to the testing session and all children were working in the classroom before the test, we do

not have measures of the teachers' compliance with our request. In future studies, researchers should either assess the teachers' compliance or include a short period of rest for the children before starting the test. Another limitation is that we only used cortisol as a measurement of stress reactivity and recovery. Including behavioral measures or measures of the sympathetic nervous system, such as cardiovascular measures and measures of salivary alpha-amylase, could render a more complete picture of young children's stress reactions to this paradigm. Finally, future research is needed to replicate our findings in larger samples. In addition, it would be interesting to determine whether the CREST is effective for younger or older age groups, and to investigate whether it can be used to study possible abnormalities in stress reactivity in clinically referred groups.

Conclusion

This study showed that our new laboratory paradigm—CREST—induced a significant increase in salivary cortisol concentration in 5- and 6-year-olds. No effects of possible confounders (child's sex, age or school, parental educational level, time of testing, sex of experimenter, and sex of judge) were observed. To our knowledge, this is the first study that shows a general cortisol response in 5- and 6-year-olds to a standardized laboratory paradigm. As such, this paradigm is a promising tool to be used in the future and might help to unravel the complex early life interactions between physiology and psychology that can lead to (psycho)pathology in adulthood.

Acknowledgments

The authors would like to thank the schools, parents, and children who participated in the study, as well as the assistants who helped collect the data.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Note

1. Removal of the outlier rendered similar significant results.

References

- Buske-Kirschbaum A, Jobst S, Wustmans A, Kirschbaum C, Rauh W, Hellhammer D. 1997. Attenuated free cortisol response to psychosocial stress in children with atopic dermatitis. *Psychom Med* 59:419–426.
- de Weerth C, Jansen J, Vos MH, Maitimu I, Lentjes EGWM. 2007. A new device for collecting saliva for cortisol determination. *Psychoneuroendocrinology* 32:1144–1148.

- Dickerson SS, Kemeny ME. 2004. Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychol Bull* 130:355–391.
- Donzella B, Gunnar MR, Krueger K, Alwin A. 2000. Cortisol and vagal tone responses to competitive challenge in preschoolers: Associations with temperament. *Dev Psychobiol* 37:209–220.
- Fairchild G, van Goozen SHM, Stollery SJ, Brown J, Gardiner J, Herbert J, Goodyer IM. 2008. Cortisol diurnal rhythm and stress reactivity in male adolescents with early-onset or adolescence-onset conduct disorder. *Biol Psychiatry* 64: 599–606.
- Gunnar M, Quevedo K. 2007. The neurobiology of stress and development. *Annu Rev Psychol* 58:145–173.
- Gunnar MR, Talge NM, Herrera A. 2009a. Stressor paradigm in developmental studies: What does and does not work to produce mean increases in salivary cortisol. *Psychoneuroendocrinology* 34:953–967.
- Gunnar MR, Wewerka S, Frenn K, Long JD, Griggs C. 2009b. Developmental changes in HPA activity over the transition to adolescence: Normative changes and associations with puberty. *Dev Psychopath* 21:69–85.
- Hasings C, Mosteller F, Tukey JW, Winsor CP. 1947. Low moments for small samples: A comparative study of order statistics. *Ann Math Stat* 18:413–426.
- Hatzinger M, Brand S, Perren S, Wyl A, Klitzing K, Holsboer-Trachsler E. 2007. Hypothalamic pituitary adrenocortical (HPA) activity in kindergarten children: Importance of gender and association with behavioral/emotional difficulties. *J Psychiatr Res* 41:861–870.
- Jansen J, Beijers R, Riksen-Walraven M, de Weerth C. 2010. Cortisol reactivity in young infants. *Psychoneuroendocrinology* 35:329–338.
- Klimes-Dougan B, Hastings PD, Granger DA, Usher BA, 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. *Dev Psychopath* 13:695–719.
- Kryski KR, Smith HJ, Sheikh HI, Singh SM, Hayden EP. 2011. Assessing stress reactivity indexed via salivary cortisol in preschool-aged children. *Psychoneuroendocrinology* 36: 1127–1136.
- Kudielka BM, Wüst S. 2010. Human models in acute and chronic stress: Assessing determinants of individual hypothalamus-pituitary-adrenal axis activity and reactivity. *Stress* 13:1–14.
- Kudielka BM, Buske-Kirschbaum A, Hellhammer DH, 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.
- Lewis M, Ramsay D. 2002. Cortisol response to embarrassment and shame. *Child Dev* 73:1034–1045.
- Luby JL, Heffelfinger A, Mrakotsky C, Brown K, Hessler M, Spitznagel E. 2003. Alterations in stress cortisol reactivity in depressed preschoolers relative to psychiatric and no disorder comparison groups. *Arch Gen Psychiatry* 60:1248–1255.
- Moons WG, Eisenberger NI, Taylor SE. 2010. Anger and fear responses to stress have different biological profiles. *Brain Behav Immun* 24:215–219.
- Petrowski K, Herold U, Joraschky P, Wittchen HU, Kirschbaum C. 2010. A striking pattern of cortisol non-responsiveness to psychosocial stress in patients with panic disorder with concurrent normal cortisol awakening responses. *Psychoneuroendocrinology* 35:414–421.
- Quas JA, Bauer A, Boyce WT. 2004. Physiological reactivity, social support, and memory in early childhood. *Child Dev* 75: 797–814.
- Schuetze P, Lopez FA, Granger DA, Eiden RD. 2008. The association between prenatal exposure to cigarettes and cortisol reactivity and regulation in 7 month old infants. *Dev Psychobiol* 50:819–834.
- Segerstrom SC, Miller GE. 2004. Psychological stress and the human immune system: A meta-analytic study of 30 years of inquiry. *Psychol Bull* 130:601–630.
- Yim IS, Quas JA, Cahill L, Hayakawa C. 2010. Children's and adult's salivary cortisol responses to an identical psychosocial laboratory stressor. *Psychoneuroendocrinology* 35:241–248.
- Von Klitzing K, Kelsay K, Emde RN. 2003. The structure of 5-year-old children's play narratives within the MacArthur Story Stem methodology. In: Emde RN, Wolf DP, Oppenheim D, editors. *Revealing the inner world of young children. The MacArthur Story Stem Battery and parent-child narratives*. New York: Oxford University Press. p 106–128.