ORIGINAL ARTICLE

Testosterone and Cortisol Levels in University Students Reflect Actual rather than Estimated Number of Wrong Answers on Written Exam

Jaroslav Flegr, Lenka Příplatová

Faculty of Science, Charles University in Prague, Viničná 7, Praha, Czech Republic.

Corresponding author: Prof. RNDr. Jaroslav Flegr, DrSc. Faculty of Science, Charles University in Prague, Viničná 7, CZ-128 44 Praha 2, Czech Republic. TEL: +(420)221951821; FAX: +(420)224919704; E-MAIL: flegr@cesnet.cz

Submitted: 2010-04-13 *Accepted: 2010-07-18* *Published online: 2010-08-27*

Key words: **steroids; aggression; competition; dominance; consciousness; subconsciousness; stress; biosocial hypothesis**

Act Nerv Super Rediviva 2010; **52**(2): 176–180 ANSR52020A12

Abstract

OBJECTIVES: After dominance-related encounters, testosterone levels increase in winners and decrease in losers. In humans, many exceptions have been described. It is possible that the complicated patterns in humans result from the methods limitations – measurement of hormone concentrations in simulated competitive events or sport instead in real-life situations.

METHODS: Here we studied changes in hormonal levels and self-estimated attractivity in real situations, namely in students after written exams.

RESULTS: We observed that the testosterone and cortisol increased or decreased in relation to the number of wrong answers on the exam. The number of wrong answers was a better predictor of the hormonal changes (increase of both testosterone and cortisol in successful, decrease in unsuccessful students) than the self-estimated number of wrong answers or a subjectively opinionated impression from the exam. On the contrary, the concentration of hormones before the exam and self-estimated attractivity were better predictors of the subjective impression from the exam than the number of wrong answers.

CONCLUSSIONS: Our results suggest that the students' subconsciousness, which directly influences the concentration of hormones, is able to objectively estimate results of an exam better than their consciousness.

INTRODUCTION

Vertebrates respond to dominance-related encounters through changes in the concentration of cortisol and testosterone depending on how the individual succeeds in the event. In animals, cortisol increases and testosterone decreases in losers while the reverse is true for victors. The biosocial hypothesis suggests that there is a feedback loop between an individual's testosterone level and his/her posterior efforts to improve or maintain social status. Winning would lead to an increase in testosterone, which in turn, stimulates competitiveness. Conversely, defeat would involve a decrease in testosterone that should reduce the possibility of engaging in new and potentially injurious encounters (Mazur 1985; Mazur & Booth 1998). In humans however, the results are less clear. For example, in judo fighters, the post-Randori competition concentrations of testosterone were higher in losers than in victors while concentrations of cortisol were higher in victors than in losers (Suay *et al* 1999). Similarly, women supported by a boyfriend in an experimental stress-induced situation reported a higher well-being during the experiment, but at the same time, expressed higher post-event concentrations of cortisol than in stranger-supported or unsupported women (Kirschbaum *et al* 1995). Furthermore, in low-motivated subjects, the increase of cortisol is typical for victors, not for losers (Wirth *et al* 2006), and the increase of testosterone in victors is typical only for highly motivated subjects (Schultheiss & Rohde 2002).

A complicated pattern of hormonal changes observed in humans can be accounted for by the limitations of the experimental methods used. The changes in hormones concentration are not measured in real-life situations, but rather in simulated competitive events, either in a psychological laboratory or during sport competitions. The release of hormones, however, is regulated with extracortical brain structures; therefore, the response of the organism in a simulated or real fight can dramatically differ. Moreover, the competitive encounters in sport are mostly accompanied with intensive physical activity, which is known to be associated with increased levels of stress hormones regardless of the results of the competition (Kuoppasalmi *et al* 1980).

To study the endocrine response of human organisms on their success or failure under natural situations, we monitored the levels of testosterone and cortisol in students before and after a written exam, and we studied the correlation between the assumed and real number of wrong answers on the test with the change in concentration of these hormones and with self-estimated attractivity.

MATERIAL AND METHODS

Design of the study

Before starting a written exam on the Methodology of Science or Evolutionary Biology, the undergraduate students of the biology programs at Faculty of Science, Charles University were asked to voluntarily and without any compensation participate in the study. The students who signed informed consents were asked to provide an initial sample of saliva immediately before the exam. At the end of the 30 minute test, the students were asked to evaluate the attractivity of the faces of foreign students screened on 4 slides – two slides with six female student photos and two slides with six male student photos – using a 7-point scale (1 – very unattractive, 7 – very attractive). They then evaluated their own attractivity using the same scale. Next, the students were asked to write the estimated number of wrong answers on their test and whether they had a positive or negative impression from the test. After approximately 10 minutes after the end of the test, the students provided a second sample of saliva. Both tests consisted of 28 questions accompanied with three wrong and one correct answers. Each question with the answer choices was shown on the screen for about 30–40 seconds (depending on the difficulty of a particular question), and if all the students had not decided on an answer within the allotted time, another 15–20 seconds was added. After that, the students wrote down a code of a presumed correct answer and a new question was screened. The tests started at 14 and 15 o'clock and 50–80 students participated in each run. More than 90% of the students consented to participate in the study but only about 50% provided saliva samples with enough material for the analyses. Five of the 293 students participated in both the Methodology and Evolutionary Biology tests and the two tests were about 3 months apart.

Radioimmunoassay tests

All hormone assays were performed at the Institute of Endocrinology, Prague with the RIA technique, using the automatic analyzer Stratec (Immunotech, Praha) and 12-channel gamma counter (Berthold, FRG) (Flegr *et al* 2008a; Flegr *et al* 2008b; Hampl *et al* 1990).

Statistics

The concentration of hormones had skewed distribution and other variables were binary or ordinal; therefore the correlations were tested using nonparametric Spearman tests. The relation between the estimated and real number of wrong answers was computed with linear regression, and differences between male and female and between below-average and an above-average students were estimated with Mann-Whitney U Test.

RESULTS

The population consisted of 214 women and 79 men. The concentrations of hormones were higher in men (Table 1); therefore we standardized these data by computing Z-scores separately for men and women. Table 1 also shows the strength and significance of the correlation of students' impression from the exam (positive or negative), the expected and actual number of wrong answers with concentration of hormones, the average attractivity of foreign students´ faces and self-estimated attractivity. Figure 1 shows the differences in the concentration of hormones in students that expected a below-average and an above-average number of wrong answers (part A), and the differences in the concentration of hormones in students that actually achieved a below-average number and an above-average number of wrong answers (part B). The results suggest that the expectation of bad results on the exam correlates with a high concentration of cortisol before the exam, while actual bad achievements were accompanied with decreased concentrations of testosterone and cortisol and good achievements with increased concentrations of both hormones after the exam.

Self-estimated attractivity correlated with a positive impression from the exams (Spearman $R = 0.153$, p=0.013). The separate analyses for men and women showed that the pattern for both sexes is similar with an exception of the absence of any correlation between impression from the exam and self-estimated attractivity

Tab. 1. Descriptive statistics and results of correlation tests.

	mean		valid N		gender		impression		expected errors		actual errors	
	women	**men**	**women**	**men**	**Z**	***p*-value**	**R**	***p*-value**	**R**	***p*-value**	**R**	***p*-value**
pre-testosterone (nmol/l)	0.169	0.467	138	58	8.926	<0.001	−0.016	0.827	−0.029	0.691	−0.038	0.593
pre-cortisol (nmol/l)	4.133	6.837	116	45	4.515	<0.001	−0.099	0.223	0.142	0.075	−0.091	0.252
post-testosterone (nmol/l)	0.143	0.455	136	58	8.839	<0.001	−0.017	0.822	−0.083	0.253	−0.246	<0.001
post-cortisol (nmol/l)	3.192	5.210	123	48	4.849	<0.001	−0.036	0.646	0.030	0.703	−0.231	0.002
diff-testosterone (nmol/l)	−0.030	−0.030	128	54	−0.179	0.858	0.088	0.248	−0.144	0.056	−0.222	0.003
diff-cortisol (nmol/l)	−1.066	−1.529	105	42	0.024	0.981	0.063	0.457	−0.061	0.466	−0.190	0.021
age (years)	21.080	21.443	213	79	1.049	0.294	−0.001	0.991	0.036	0.541	0.094	0.109
impression (0 –bad, 1 good)	0.215	0.346	209	78	1.742	0.117			−0.686	<0.001	−0.288	<0.001
expected errors	11.729	10.608	210	79	−1.731	0.083	−0.686	<0.001			0.471	<0.001
actual errors	9.724	9.759	214	79	0.213	0.831	−0.288	<0.001	0.471	<0.001		
allo-attractivity (1 unattract. 7 attract.)	3.381	3.077	214	77	−2.856	0.004	0.032	0.597	0.023	0.704	−0.011	0.847
self-attractivity (1 unattract. 7 attract.)	4.550	4.360	200	75	−1.625	0.104	0.153	0.013	−0.096	0.116	0.104	0.087

First six columns show descriptive statistics for population and results of testing difference between men and women with Mann-Whitney U test. The last six columns show results of Spearman correlation test; for hormones the tests were performed on Z-scores computed separately for men and women.

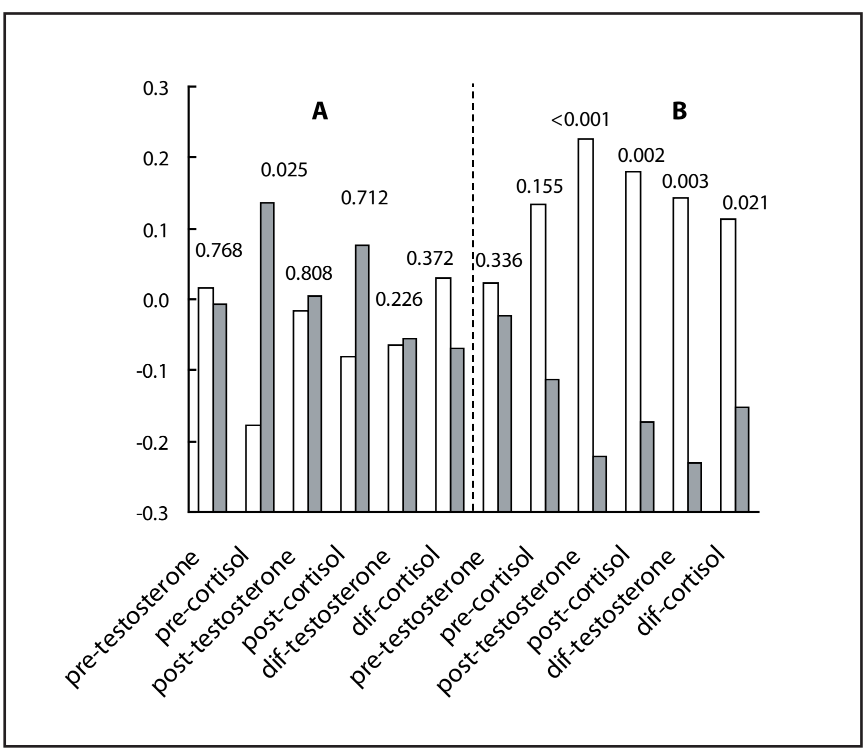

Fig. 1. Concentration of testosterone and cortisol before and after exam. The concentrations (y-axis) were expressed as Z-scores. Empty columns: concentrations for students that expected below-average number of wrong answers (part A) or for students that achieved below-average number of wrong answers (part B), gray columns: concentrations for students that expected above-average number of wrong answers (part A) or achieved above-average number of wrong answers (part B). Differences in concentration were computed as Z-score of post-exam concentration – Z-score of pre-exam concentration, i.e. the high values indicate an increase of concentration of particular hormone during the exam. The results (*p*-values) of Mann-Whitney tests are shown above particular columns.

in men; however, both the strength of the correlations and significance were always higher for women than men.

The estimated number of wrong answers strongly correlated with the actual number of wrong answers (beta = 0.477, R^2 = 0.22, p<0.001; Figure 1). The residuals of this correlation (reflecting self-confidence or optimism of the subjects) correlated positively with self-estimated attractivity (Spearman R = 0.186, p=0.002) and negatively with the level of testosterone after the exam (Spearman R = –0.210, p=0.004), the level of cortisol after the exam (Spearman R = –0.281, p<0.001), and the change in testosterone (Spearman R = –0.188, p=0.012) and cortisol levels (Spearman R = –0.218, p=0.009) during the exam. Separate analyses for men and women showed a similar pattern, however, no correlations were significant for males.

DISCUSSION

The concentration of testosterone and cortisol changed in relation to the number of wrong answers on an exam. The actual number of wrong answers was a better predictor of the change in hormones (increase in successful, decrease in unsuccessful students) than the self-estimated number of wrong answers or a subjective positive or negative impression from the exam. On the contrary, the concentration of hormones before the exam was a better predictor of the subjective positive or negative impressions from the exam than the actual number of wrong answers on the test. Self-estimated attractivity of females but not males correlated negatively with the subjective positive impression from the exam; however, there was no correlation between the self-estimated attractivity and the actual number of wrong answers on the exam.

The relation between success on the exam and concentration of hormones was stronger for females than males. This was not caused by the lower number of men in the experimental set since the statistical significances as well as the absolute values of Spearman R were lower for men than for women. We cannot decide whether the concentration of hormones more closely reflects a response to success in women than in men, or whether men were less concerned about their results in the exams than women were. The latter possibility was supported by the observed absence of any relation between self-attributed attractivity and results of the exam in the men.

Similarly, reasons for the stronger relation between the concentrations of hormones and the actual rather than estimated number of wrong answers cannot be definitively determined. We can speculate that the students' subconscious, which directly influences the concentration of hormones, is able to objectively estimate the results of the exam better than their consciousness. It is also possible that the weak students consider, for example, 10 wrong answers to be a good result while good students consider the same number of wrong answers to be a failure. A better correlation between the concentration of hormones and the number of actual rather than estimated errors can be expected if the hormone concentration is influenced by the real number of failures (e.g. inabilities to decide the correct response for a particular question) rather than the rational reinterpretation of what is a success and what is a failure. It is also possible that the weaker association between the estimated number of wrong responses and the concentration of hormones was caused by systematic bias, e.g. the inability of weak students to correctly estimate the number of wrong answers.

The major shortcoming of the study was the relatively mild (non-stressing) conditions of the exams. For ethical reasons, we used the same benevolent and anti-stress conditions as in normal exams undertaken in previous years. The students voluntarily participated in these exams, and in the cases where they were not satisfied with their results, they had the option of completing an oral exam where the result of their written exam was not taken into consideration. Moreover, the administration of the test was set up to minimize the pressure and stress of time; the next question was always screened after all students had decided which answer they considered to be correct. Clearer results could probably be obtained if participation in the exam was compulsory, the students did not have the option to refute the results of the exam, and the questions were administered under time-pressure conditions.

The present study confirmed that the concentration of testosterone and cortisol increases or decreases in response to natural psychological stressors depending on individual achievement. The subjects that were successful increased their levels of testosterone and cortisol while unsuccessful subjects expressed a decrease in concentration of these hormones. It is not clear whether an increased level of cortisol, which is known to mobilize energy to cope with a stressing event (Kuoppasalmi *et al* 1980), was partly responsible for better achievements in successful students, or whether the concentration increased in response to their good achievement as is usually suggested for testosterone (Mazur 1985; Mazur & Booth 1998). Usually, but not always, testosterone increases and cortisol decreases after victory, and testosterone decreases after defeat in humans (Salvador 2005). However, this typical response can be reversed in various situations and depends on many variables, including the non-conscious motivation of subjects (Schultheiss & Rohde 2002; Wirth *et al* 2006), the social situation (Kirschbaum *et al* 1995; Newman *et al* 2005), or the experience and manner in which the subject copes with the stress (Salvador 2005; Suay *et al* 1999).

Our results suggest that the complicated pattern of the hormonal response observed in humans is not an artifact of methods used in human research, namely of the fact that the hormonal responses are usually studied in simulated rather than real competitive events.

The most interesting result was the observation that an increase/decrease of hormones correlated with the actual rather then with the expected number of wrong responses or with subjective impression from the exams. It has already been reported that the extracortical parts of the brain regulate the secretion of steroid hormones (Parmigiani *et al* 2006). It is not very surprising that the extracortical rather than the cortical structures can better "estimate" success or failure in social interactions, e.g. in competitions for sexual partners. However, it seems rather surprising that the students' subconsciousness can also better "estimate" how a subject succeeds on a written university exam than their consciousness.

ACKNOWLEDGMENTS

The research was supported by MSMT, project 0021620828.

REFERENCES

1 Flegr J, Lindová J, Pivoňková V, Havlíček J (2008a). Brief Communication: latent toxoplasmosis and salivary testosterone concentration-important confounding factors in second to fourth digit ratio studies. Am J Phys Anthropol. **137:** 479–484.
2 Flegr J, Novotná M, Lindová J, Havlíček J (2008b). Neurophysiological effect of the Rh factor. Protective role of the RhD molecule against *Toxoplasma*-induced impairment of reaction times in women. Neuroendocrinology Letters. **29:** 475–481.
3 Hampl R, Foretova L, Sulcova J, Starka L (1990). Daily profiles of salivary cortisol in hydrocortisone treated children with congenital adrenal-hyperplasia. Eur J Pediatr. **149:** 232–234.
4 Kirschbaum C, Klauer T, Filipp SH, Hellhammer DH (1995). Sex-specific effects of social support on cortisol and subjective responses to acute psychological stress. Psychosom Med. **57:** 23–31.
5 Kuoppasalmi K, Naveri H, Harkonen M, Adlercreutz H (1980). Plasma-cortisol, androstenedione, testosterone and luteinizing-hormone in running exercise of different intensities. Scand J Clinic Lab Invest. **40:** 403–409.
6 Mazur A (1985). A biosocial model of status in face-to-face primate groups. Soc Forc. **64:** 377–402.
7 Mazur A, Booth A (1998). Testosterone and dominance in men. Behav Brain Sci. **21:** 353–363.
8 Newman ML, Sellers JG, Josephs RA (2005). Testosterone, cognition, and social status. Horm Behav. **47:** 205–211.
9 Parmigiani S, Bartolomucci A, Palanza P, Galli P, Rizzi N, Brain PF, Volpi R (2006). In judo, Randori (free fight) and Kata (highly ritualized fight) differentially change plasma cortisol, testosterone, and interleukin levels in male participants. Aggresiv Behav. **32:** 481–489.
10 Salvador A (2005). Coping with competitive situations in humans. Neurosci Biobehav Rev. **29:** 195–205.
11 Schultheiss OC, Rohde W (2002). Implicit power motivation predicts men's testosterone changes and implicit learning in a contest situation. Horm Behav. **41:** 195–202.
12 Suay F, Salvador A, Gonzalez-Bono E, Sanchis C, Martinez M, Martinez-Sanchis S, Simon V, Montoro JB (1999). Effects of competition and its outcome on serum testosterone, cortisol and prolactin. Psychoneuroendocrinology. **24:** 551–566.
13 Wirth MM, Welsh KM, Schultheiss OC (2006). Salivary cortisol changes in humans after winning or losing a dominance contest depend on implicit power motivation. Horm Behav. **49:** 346–352.