Testing the Relationship Between Levels of Endogenous Testosterone and Physiological Responses to Facial Expressions in Men: An Experiment Conducted by Students in an Undergraduate Behavioral Neuroscience Class

Richmond R. Thompson & Kirsten George
Psychology Department and Neuroscience Program, Bowdoin College, Brunswick, ME 04011

To determine if endogenous testosterone (T) is related to physiological responses to aggressive stimuli in human males, students in a behavioral neuroscience laboratory class conducted an experiment that determined if levels of salivary T in adult males are correlated with autonomic and/or somatic responses to angry facial expressions. Each student collected a saliva sample from one subject and, within 30 minutes of collecting the sample, measured heart rate (HR), skin conductance (SC), and corrugator supercili electromyographic (EMG) responses to emotionally neutral, happy, and angry male facial expressions. Salivary T levels were analyzed by an enzyme-linked immunoassay. A significant, positive correlation was found between levels of salivary T and HR responses to angry and happy, but not neutral, male facial expressions. This laboratory experience not only provided students with the opportunity to design and conduct a scientific experiment, but it also generated preliminary data suggesting that levels of salivary T collected within 30 minutes of testing are related to autonomic responses to emotional social stimuli in humans. If verified by future experiments, this finding would be consistent with the hypothesis that fluctuations in circulating T might influence ongoing social behavior in human males by rapidly modulating autonomic responses to emotional social stimuli. The potential significance of such a general mechanism for the regulation of aggressive behavior is discussed.

Keywords: testosterone, aggression, human, angry, autonomic

Testosterone (T) is known to be involved in the regulation of aggressive behavior in numerous species from diverse vertebrate groups, which indicates a strong conservation of this behavioral function during evolution. However, the relationship between T and aggression in humans has remained controversial (Rubinow and Schmidt, 1996; Archer, 1991; Albert et al., 1993). Positive correlations have been reported between levels of T and self-reported levels of aggression (Albert et al., 1993) and the levels of aggression reported by others who interact with those individuals (Gerra et al., 1996). Similarly, positive correlations have been found between T levels and the degree of violence associated with the crimes for which males are imprisoned (Dabbs et al., 1987) and the records of aggressive behavior in prison in female inmates (Dabbs and Hargrove, 1997). T levels have also been reported to be positively associated with the likelihood of responding vigorously to provocation and threat (Olweus et al., 1988), with developmental increases in aggression during puberty (Olweus et al., 1980), and with the likelihood of acting aggressively during non-play social interactions in preschool children (Sanchez-Martin et al., 2000). On the other hand, numerous studies have reported no correlations between T levels and various parameters of aggression, including violence in prison in male inmates (Kreuz and Rose, 1972; Mattson et al., 1980) and levels of aggressive behavior during puberty (Susman et al., 1987; see Albert et al., 1993, for a more complete review of studies reporting negative data).

In support of the hypothesis that T can increase aggressive behavior in humans, treatments that increase T levels have been shown to increase self-reported levels of aggression and aggressive impulses in hypogonadal adolescents (Finkelstein et al., 1997) and to increase "aggressive" responses in normal men in a competitive points subtraction game (Kouri et al., 1995). Again, though, evidence contradicting this hypothesis also exists; elevating T levels failed to increase self-reported anger or that observed by a close partner in normal men (Tricker et al., 1996) or self-reported aggression in hypogonadal or eugonadal men (O'Connor et al., 2002). In addition, decreasing T levels with a GnRH antagonist in normal men did not result in decreased measures of self-reported aggressiveness in a study by Bagatell et al. (1994).

One potential reason for the inconsistencies that have been reported regarding the relationship between T and aggressive behavior in humans lies in the difficulties associated controlling the complex social contexts that constrain the behavioral manifestations that any hormonal influences on processes related to aggression could have. However, studies have begun to elucidate the specific physiological and cognitive processes affected by androgens that may, in permissive social contexts, be able to influence the expression of aggressive behavior in humans. For example, levels of T are positively related to measures of impulsivity in women (Bjork et al., 2001) and to selective attention toward angry facial stimuli in men and women (van Honk et al., 1999), either of which could, in the appropriate conditions, increase the likelihood that an individual will participate in aggressive behaviors. More directly, testosterone administration in women has been shown to accelerate cardiac responses to...
angry/threatening facial expressions (van Honk et al., 2001). A selective enhancement of autonomic/emotional responsiveness to threatening stimuli could make some individuals more likely to engage in aggressive behavior when a threat emerges or increase the intensity of an aggressive response once initiated.

The study by van Honk et al. (2001) showing an effect of T administration on autonomic reactivity to threatening facial stimuli in women is particularly intriguing because the injections were given only 30 minutes before testing, which suggests that T's effects on this process may be mediated by rapid membrane effects of the sex steroid rather than by a slow, genomic mechanism. Such a mechanism would make it possible for fluctuations in circulating T, potentially in response to aggressive stimuli, to feed back to modulate responses toward those stimuli. In contrast, the relationship between levels of testosterone and selective attention toward angry facial stimuli were only apparent for testosterone samples collected six hours prior to the test (van Honk et al., 1999). However, the physiological relevance of the rapid effect of T on heart rate responses to angry facial stimuli remains undetermined, as does whether or not a similar mechanism is operative in men.

To address these questions and to provide students in a behavioral neuroscience laboratory course with experience in experimental design, data collection, statistical analysis and paper writing, students in Techniques in Behavioral Neuroscience at Bowdoin College were instructed to design an experiment that would determine if endogenous T levels are related to physiological responses to aggressive stimuli in human males. Students were given appropriate background readings (Archer, 1991; Bjork et al., 2001; Rubinow and Schmidt, 1996; van Honk et al., 1999; van Honk et al., 2001; and selected chapters from Nelson, 2000) and lectures that covered mechanisms of steroid hormone actions and the behavioral effects of T in humans and in animal models, as well as various methodological papers that exposed them to some of the techniques that have been used to answer similar questions. Together, the students and instructor designed the protocol for the study, in which students tested if levels of salivary T from samples collected 30 minutes prior to testing are correlated with autonomic and/or somatic responses to emotional facial expressions in human males. Males were chosen for this initial experiment because there are more studies linking T to aggressive behavior in males than in females. Based on past findings that T injections can increase HR responses to angry facial stimuli (van Honk et al., 2001), our primary hypothesis was that salivary T levels would be positively correlated with HR accelerations to angry facial expressions. Two secondary hypothesis were that T levels would be positively correlated with skin conductance (SC) responses, another measure of autonomic activity, and/or with corrugator supercilii electromyographic (EMG) responses to the same stimuli. The corrugator supercilii is a muscle group just above the brow in the face whose activity is associated with agonistic communication (Jancke, 1996).

**MATERIALS AND METHODS**

Thirteen male students in an introductory psychology class at Bowdoin College (18-22 years old) were used as subjects. Subjects received credit for research participation in their introductory psychology class (students in Psych 101 at Bowdoin have the option of participating as subjects in ongoing research projects or writing papers about research as part of the requirement for that course). Each subject read and signed a consent form approved by the research oversight committee at Bowdoin College before beginning the experiment. Each student experimenter in the class tested one subject. Immediately after signing the form, subjects were instructed to wash their hands and arms thoroughly with anti-bacterial soap and water and to dry them on paper towels. Three to four ml of saliva were then collected from individual subjects by instructing them to spit through a straw into a plastic centrifuge tube. The experimenter wore protective eyeglasses and gloves during the saliva collection procedure and thoroughly washed her/his hands after collection was complete. All saliva was disposed in biohazard receptacles. Vials were kept on ice during the experimental procedure and then stored at 80°C until further analysis.

While saliva was being collected, electrodes were attached to collect SC and HR, two measures of autonomic activity, and corrugator supercilii EMGs, a somatic response related to agonistic communication (Jancke, 1996). SC was measured directly by an isolated skin conductance coupler, using a constant of 0.5 V, through a pair of 9 mm Ag/AgCl electrodes filled with an isotonic paste and placed within 5 mm of one another on the hypothenar surface of the non-dominant hand (the smooth area along the bottom of the palm). Placement sites for EMG and HR (see below) were slightly abraded with a rough pad and then rinsed with rubbing alcohol to reduce impedance to below 5 KΩ. For EMG, 4 mm Ag/AgCl surface electrodes were filled with electrolyte paste and placed on the left corrugator supercilii, just above the brow over one eye, according to the placement diagrams of Fridlund and Cacioppo (1986), and one ground electrode was placed in the center of the forehead. The EMG signal was amplified by an isolated bioamplifier with bandpass filter (90 - 1000 Hz); a 200 ms time constant was used for integration by a multi-function integrator. One electrode was placed on each forearm to collect HR, which was recorded from standard limb electrocardiogram leads connecting to an isolated bioamplifier with a bandpass filter through which interbeat intervals were converted to heart rate. All physiologic analog signals were digitized by a general purpose lablinic port and sampled using a Dell Otiplex GX1 Pentium III computer at the rate of 1000 KHz.

Once all electrodes were attached and within 30 minutes of saliva collection, the stimulus presentation began. Each subject was instructed to fix his gaze in the center of a blank computer screen. Every 20-30 sec, an image of a male face was presented for eight sec. The facial expressions varied in their emotional content: eight were emotionally neutral, eight were happy, and eight were
angry. One of each expression from eight different models was used. Happy faces were included as a control for emotional valence so that we would be able to determine if any relationships between T and physiological responses to angry facial stimuli are specific for negative, threatening emotional stimuli. Faces with different emotional expressions were presented in random order. The emotional content for all facial stimuli, which were selected from Ekman and Friesen's Pictures of Facial Affect (1976), has been validated. All faces in this stimulus set are white, as were all subjects tested. Stimulus presentation, baseline SC, HR, and corrugator supercilii EMGs, and time-locked, stimulus induced responses were captured by a software package (Modified Emotional Stroop) produced by Dan Sussman of Coulbourne Electronics. Baseline responses were recorded during the five sec immediately prior to each image, and stimulus induced responses were captured during the eight sec while the image was presented. Average baseline responses were subtracted from the peak responses during the eight sec while the stimulus was presented. Peak HR responses are typically used as a measure of HR acceleration: (Lang et al., 1993).

To quantify salivary testosterone, an enzyme-linked immunoassay (EIA) kit for testosterone was purchased from Cayman Chemicals. Each testosterone sample was briefly centrifuged and 100 ul of the clear supernatant was diluted 1/10 and 1/1000 in buffer. Duplicate reactions for each sample at both dilutions were run by incubating 50 µl of each in wells with a constant amount of T-antibody and T-tracer (T linked to acetylcholinesterase) according to the Cayman protocol. A standard curve was generated by incubating the same amounts of T antibody and tracer with a serial dilution of known amounts of T. All wells were incubated with a substrate (Ellman's reagent) that, in the presence of acetylcholinesterase, produces a precipitate that absorbs light at 412 nm, and the absorbance was quantified with a plate reader.

The technician for the class performed the pipetting associated with the actual assay because the plate reader was only available for use outside of regular class hours, but the actual readings from the assay were given to each individual student for analysis. Also, it would be prohibitively expensive to have each student run his/her own plate and scientifically unsound to have each student do part of a single plate, as the variation would increase dramatically. Each student did receive and use the raw absorbance readings to generate the equation for the standard curve using linear regression, and then used that equation to determine the amounts of T in each subject's sample. Data from only the 1/10 dilution gave readings that were within the reliable detection limits of the assay (not falling below 20% or above 80% maximal binding). Each student was also given the raw data from the physiological recordings obtained for each subject.

**Statistics**

Although a basic statistics course was a prerequisite for this course, most students did not have any experience with the application of statistics to actual data sets. The class therefore offered an excellent opportunity to discuss issues related to setting up appropriate statistical tests in conjunction with the development of the actual experimental protocol. In particular, we focused on the importance of organizing the experiment so as to maximize the power of the analysis, especially since we knew we would not be able to generate sample sizes as large as would be optimal, by making a few, very specific predictions. In this way the loss of power that accompanies the application of methods that reduce the likelihood of Type I errors when multiple tests are performed could be avoided. In that context, we discussed the potential application of Bonferroni corrections to our data, but because we had only one primary, a priori prediction (that T and HR responses to angry expressions would be positively correlated) and two secondary, a priori predictions (that T and EMG or SC responses to angry expressions would be positively correlated), we decided not to apply them in this pilot study, recognizing that this decision could limit the strength of our interpretation and make it necessary for us to consider any marginal findings as completely preliminary.

According to our specific hypotheses, each student determined if there was a correlation between levels of salivary testosterone and physiological responses to the angry facial stimuli using Pearson correlations (three tests total). Alpha levels were set at p < 0.05 for statistical significance. If a significant relationship was found for any of the dependent variables, students then tested the relationship between T levels and responses to the neutral and happy expressions for that variable. T levels for one subject were not detected and so all dependent variables for that subject were excluded, and HR responses for one subject were not recorded due to a faulty electrode attachment and so were excluded. Also, independent of our hypothesis about the relationship of T to physiological responses to particular classes of emotional stimuli, students conducted a repeated measures, one-way ANOVA to determine if the different classes of stimuli differentially affected the dependent measures recorded. An additional, post-hoc ANOVA for HR data was also run using T as a covariate.

**RESULTS**

The data presented are those analyzed by one student, but the same findings were reported by other students in the class. The data set provided opportunities to discuss many important issues related to that application of statistical methods to actual, raw data. In particular, we discussed the assumptions of parametric statistical tests, especially the importance that the dependent measures are normally distributed. Because the HR responses were positively skewed and because many values were less than 1, 1 was added to the raw values and the square root of the resulting values were used for statistical testing, as suggested for such data sets by Snedecor and Cochran (1989). The correlation between levels of salivary T and transformed HR acceleration responses to the angry facial expressions was significant (r = 0.64, p = 0.032, two-tailed, see Fig 1A).
Planned, follow-up tests of the specificity of the relationship revealed there was not a significant relationship between T and HR accelerations to neutral facial expressions (r = -0.1, p = 0.75; see Fig 1B), but that there was a significant relationship between T and HR accelerations to the happy facial stimuli (r = 0.7, p = 0.016; see Fig 1C). There were no significant correlations between T levels and SC or EMG responses to the angry facial expressions. One-way ANOVAs indicated that neither the happy nor the angry facial expressions induced changes in SC, HR or corrugator supercilii EMGs. A post-hoc ANOVA for HR was done after the significant associations between T and HR responses to angry and happy expressions were found using T as a covariate. The interaction between T and stimulus type was, as predicted, nearly significant (R (2,18) = 3.43, p = 0.055), and the main effect of stimulus type went from F (2, 22) = 0.58, p = 0.65 in the analysis without T as a covariate to F (2, 18) = 2.0, p = 0.16 in the analysis with T as a covariate. The intraassay coefficient of variation, calculated as the average coefficient of variation across the samples, was 21%. However, we suggest that others attempting this protocol run independent samples for calculating the intraassay coefficient of variation.

**DISCUSSION**

This experiment provided advanced students in behavioral neuroscience with the opportunity to participate directly in the scientific process. They designed their own experiment, collected endocrine and psychophysiological data, dealt with the complexities of data analysis, and had to discuss the implications and limitations of their findings in laboratory reports written in the format of a journal article. Furthermore, in the process they generated preliminary data suggesting that levels of salivary T, which reflect levels of free T in the body, may be immediately (at least within 30 min) related to HR accelerations in response to emotional social stimuli. This would be in contrast to the relationship between levels of T and cognitive responses to aggressive stimuli found in humans in a previous study, as correlations between T and the attention paid to angry/threatening stimuli are only apparent when samples are collected at least six hours prior to behavioral testing (van Honk et al., 1999). Thus, if future studies verify the relationship suggested by this study, then T may be related to cognitive and autonomic responses to social stimuli via dissociable mechanisms in humans.

The relationship between T and aggressive behavior in humans has been controversial, at least in part due to the difficulties associated with measuring aggressive behavior in human subjects. Because steroid hormones, including T, do not simply activate fixed motor patterns but rather modify the probability that a behavior will occur in the presence of the appropriate releasing stimuli, any potential influences of T upon behavior will only be observable in particular social contexts, which are difficult to control in human studies. However, it is possible to elucidate the underlying physiological mechanisms associated with T that may, in such contexts, be able to affect the expression of aggressive behavior. T has been shown to produce rapid effects on HR responses to threatening facial stimuli in humans (van Honk et al., 2001), and our class data suggest that variations in levels of endogenous T may be immediately related to HR responses to such stimuli. It is therefore possible that T

![Figure 1](image-url)
could increase the likelihood or intensity of aggressive behavior in permissive social contexts and/or in certain individuals by enhancing emotional responses to threatening social stimuli. Such a mechanism could, in part, explain the positive relationships that have been found between $T$ and various parameters of aggression in humans (see introduction). It could also explain the negative results found in other studies, as positive relationships may only be apparent if $T$ samples are collected at or near the time when aggressive responses typically occur.

On the other hand, these preliminary findings are not consistent with van Honk et al.’s (2001) finding that $T$ injections did not immediately increase HR responses to happy facial expressions in women. Interestingly, men, but not women, exhibit increased HR accelerations to happy faces relative to angry ones (Jonsson and Sonnby-Borgström, 2003), so there may be differences in the processing of emotional social stimuli in men and women and therefore differences in the relationship of $T$ to that processing. It would obviously be interesting to repeat this study in women in a future class experiment. The three classes of facial stimuli did not evoke statistically different HR responses (see further discussion below), but responses were highest towards happy, next highest towards angry, and lowest towards neutral facial expressions, as would be expected according to the data of Jonsson and Sonnby-Borgström (2003). If the positive correlation between $T$ and HR accelerations to happy faces can be verified in men, it would indicate that endogenous $T$ is not selectively related to stimulus-specific physiological responses associated with aggression, but rather that it may generally influence the amplitude of a variety of emotional responses. The behavioral output associated with such a mechanism would then be completely context specific. In fact, there is evidence that $T$ levels are not only related to aggression in humans, but to positive affect as well (Booth et al., 1989; Alexander et al., 1997; O’Connor et al., 2002).

As mentioned, there was not a main effect of stimulus type on HR responses, but we do not believe this negates the importance of potential relationships between $T$ levels and HR responses within particular stimulus classes. It is even possible that $T$-dependent variation within the emotional classes worked against the ability to detect a main effect of stimulus type, as lower HR responses to the emotional stimuli in individuals with low $T$ decreased the magnitude of the difference between these stimuli and those in response to the neutral expressions. In fact, specific relationships between $T$ and HR responses to emotional stimulus suggests that the ANOVA should have incorporated $T$ as a covariate. When this was done, the $T\times$ stimulus interaction was nearly significant, as expected based on the specific correlations between $T$ and responses to angry and happy expressions. Also, the main effect of stimulus type was strengthened, though it still did not reach statistical significance, therefore suggesting that $T$’s relationship to responses to emotional stimuli is stronger than $T$-independent variation between stimulus types.

We did not observe any significant relationships between salivary $T$ levels and responses to angry facial expressions in SC, another index of autonomic activity, or corrugator supercilii EMGs, a somatic response related to agonistic communication. Thus, $T$ levels may be selectively related to HR responses to emotional social stimuli. On the other hand, there are several reasons to interpret the meaning of the negative results in this study with extreme caution. First, there was undoubtedly increased variance related to the fact that a different experimenter tested each subject, which would decrease the likelihood of observing a significant relationship even if one does exist. Second, the inability of the emotional stimuli to provoke robust physiological responses relative the emotionally neutral control expressions suggests that the threatening facial expressions in the stimulus set were not particularly arousing, and so it remains possible that relationships between $T$ and SC and/or corrugator supercilii EMGs might be revealed if more emotionally evocative social stimuli are used in future studies.

In conclusion, our class experiment has provided preliminary data suggesting that endogenous levels of $T$ may be immediately related to autonomic responses to emotional social stimuli in human males. If verified by future experiments, this would strengthen the argument that fluctuations in endogenous $T$ may be capable of modulating physiological responses to social stimuli in humans. Furthermore, our findings suggest that endogenous $T$ levels in men may not be specifically related to emotional responses to threatening stimuli, but rather that they may be part of a more general mechanism that could influence responses to positive as well as negative emotional social stimuli. This experiment used basic human physiological recording equipment, a relatively simple software package, and a non-radioactive hormone assay system, all of which made the experiment possible to complete in an undergraduate classroom setting. Although the available facilities and specific expertise of the instructor will obviously vary at different undergraduate institutions, we believe the basic framework used in this class, which involves having students design and execute novel, scientific experiments as a group, is one that can be easily adapted to fit a wide range of neuroscience laboratory courses.

REFERENCES


Received February 28, 2003; revised May 01 2003; accepted May 12, 2003

This work was supported by the administration at Bowdoin College. The authors wish to thank Joyce Thompson for exceptional technical assistance and the following students, all of whom collected data as part of their laboratory course work: Shawn Pelletier, Sarah Abbott, Audrey Gray, Jessica Stern, Elena Timoshkin, Evelyne Tsang, Joseph Sansone, Cabul Mehta, Patrick Burns, Tyler Dunphy, Elisabeth Gruenberg, Thurston Riday.

Address correspondence to: Dr Richmond Thompson, Psychology Department, Bowdoin College, Brunswick, ME 04011. rtthomps@bowdoin.edu.