Testosterone and Aggressive Behavior in Juvenile Delinquents: The Mediating Role of Hostile Attentional Bias and the Moderating Role of Cortisol

Authors: Ren Zhihong, Zhao Ziyi, Yu Xianglian, Zhao Chunxiao, Zhang Lin, Lin Yuzhong, Zhang Wei, Zhang Wei

Date: 2020-07-03T00:00:00+00:00

Abstract

To examine the relationship between testosterone and aggressive behavior in juvenile offenders, this study investigates the mediating role of hostile attentional bias and the moderating role of cortisol by integrating the biohormonal perspective and social cognitive perspective of aggressive behavior. Hormone levels, hostile attentional bias, and aggressive behavior were measured in 84 juvenile offenders. The results showed that: (1) Hostile attentional bias (attentional instability, attentional avoidance) fully mediates the effect of testosterone on aggressive behavior; (2) Cortisol moderates the relationship between testosterone and hostile attentional bias, and the mediating effect of hostile attentional bias is significant only at high cortisol levels. Based on the mediating and moderating mechanisms through which biohormones influence aggressive behavior, interventions for aggressive and violent behavior in juvenile offenders may be attempted by increasing attentional avoidance of hostile stimuli and elevating cortisol levels.

Full Text

Testosterone and Aggressive Behavior in Juvenile Offenders: The Mediating Role of Hostile Attention Bias and the Moderating Role of Cortisol

REN Zhihong[1], ZHAO Ziyi[1], YU Xianglian[2], ZHAO Chunxiao[1], ZHANG Lin[1], LIN Yuzhong[3], ZHANG Wei[1]

[1]Key Laboratory of Adolescent Cyberpsychology and Behavior (CCNU), Ministry of Education; School of Psychology, Central China Normal University;

Key Laboratory of Human Development and Mental Health of Hubei Province, Wuhan 430079, China
[2]Department of Education, Jianghan University, Wuhan 430056, China
[3]School of Humanities and Social Sciences, Fuzhou University, Fuzhou 350108, China

Abstract

This study examined the relationship between testosterone and aggressive behavior in juvenile offenders by integrating biological hormone and social-cognitive perspectives. Specifically, we investigated the mediating role of hostile attention bias and the moderating role of cortisol. Eighty-four juvenile offenders completed measures of hormone levels, hostile attention bias, and aggressive behavior. Results indicated that: (1) Hostile attention bias (attentional instability, attentional avoidance) fully mediated the effect of testosterone on aggressive behavior; (2) Cortisol moderated the relationship between testosterone and hostile attention bias, such that the mediating effect was significant only at high cortisol levels. Based on these mediating and moderating mechanisms of biological hormones influencing aggressive behavior, interventions for juvenile offenders' violent behavior may target increasing attentional avoidance of hostile stimuli and elevating cortisol levels.

Keywords: aggressive behavior; testosterone; cortisol; attention bias; juvenile offender

1. Introduction

Juvenile crime represents a serious social problem that has attracted widespread public concern due to the unique characteristics of juvenile offenders, the complexity of offense types, and the extensive negative consequences. In recent years, juvenile crime in China has become increasingly severe, showing trends toward younger ages and greater violence (Yang, Xiao, Lai, & Wu, 2018). Among various forms of juvenile crime, aggressive violence constitutes the primary type (Lu et al., 2018), and its development is substantially influenced by biological predispositions (Rosell & Siever, 2015). On one hand, biological factors affect aggressive behavior primarily through hormonal regulation (Fernàndez-Castillo & Cormand, 2016), with testosterone and cortisol playing particularly important roles (Montoya, Terburg, Bos, & Van Honk, 2012). On the other hand, the effects of biological factors also depend on cognitive processing of environmental information, where selective encoding and attention bias may play crucial roles (Huesmann, 2018).

Given the significant influence of both hormones and social-cognitive processes on aggressive behavior, researchers have attempted to understand individual aggression from biological hormone perspectives (Mehta & Josephs, 2010) and social information processing perspectives (Crick & Dodge, 1994; Huesmann, 2018; Wilkowski & Robinson, 2010). However, few studies have considered the close connections between these two perspectives and their joint influence on aggressive behavior. From an integrative perspective, it is necessary to explore the underlying mechanisms and influencing factors of aggressive behavior in juvenile offenders, which is important for establishing effective monitoring and intervention measures to reduce aggression levels.

1.1 Biological Hormone Perspective: Testosterone and Aggressive Behavior

For group animals including humans, aggressive behavior serves as an important means to gain control and dominance, and biological hormones constitute a key factor influencing aggression levels (Fernàndez-Castillo & Cormand, 2016). Testosterone, produced by the HPG (hypothalamic-pituitary-gonadal) axis, is particularly closely linked to aggressive behavior (Montoya et al., 2012). As an endocrine hormone that motivates individuals to pursue higher social status, testosterone is often associated with aggressive behaviors that demonstrate control and dominance (Eisenegger, Haushofer, & Fehr, 2011), and can exert automatic, unconscious influences on aggression by modulating amygdala reactivity to threatening stimuli (Terburg & van Honk, 2013).

Early studies on adult criminals (Dabbs, Carr, Frady, & Riad, 1995) and juvenile offenders (Brooks & Reddon, 1996) reported positive predictive effects of testosterone on aggressive behavior. However, other studies have found non-significant correlations between testosterone and aggression (Panagiotidis et al., 2017). These inconsistent findings suggest that the influence of biological hormones on behavior is also affected by other psychosocial factors (Huesmann, 2018). In addition to biological hormones, human behavior more actively responds to the environment, with social cue information processing playing an important role in this process (Crick & Dodge, 1994).

1.2 Social Information Processing Perspective: Attention Bias and Aggressive Behavior

The social information processing perspective posits that aggressive behavior is closely related to biased cognitive processing patterns (Crick & Dodge, 1994; Huesmann, 2018; Wilkowski & Robinson, 2010), particularly attention bias toward hostile stimuli (Miller & Johnston, 2019). Social information processing theory suggests that selective attention filters information and can influence subsequent cue interpretation and behavioral responses in a bottom-up manner (Crick & Dodge, 1994). On one hand, individuals who more readily notice hostile social cues tend to perceive more threats, experience greater negative emotion arousal, and consequently show more aggressive behavior (Crick & Dodge, 1994; Wilkowski & Robinson, 2010). On the other hand, selective encoding of hostile stimuli also makes individuals more likely to interpret others' intentions as hostile, forming hostile cognitive beliefs and information processing schemas that ultimately lead to general aggressive tendencies (Miller & Johnston, 2019). Consistent with this view, empirical studies on violent offenders (Domes, Mense, Vohs, & Habermeyer, 2013), aggressive adolescents (Laue et al., 2018), and meta-analytic research (Manning, 2019) have all demonstrated positive associations between attention bias toward hostile stimuli and aggressive behavior.

Integrating these two perspectives, it is evident that human aggressive behavior is not only a product of specific biological hormones but also reflects the influence of higher-level cognitive processing, with the two being closely interconnected. As individuals develop, the influence of hormones on aggressive behavior increasingly depends on cognitive control and information processing (Archer, 2006). Combining these two perspectives may help us investigate the influencing factors, internal mechanisms, and individual differences in aggression from a more systematic viewpoint.

1.3 Integration of Biological Hormone and Social Information Processing Perspectives: A Moderated Mediation Model

Based on the connections between biological hormones and social information processing and their effects on aggressive behavior, we propose a moderated mediation model: while considering the influence of biological hormones, we treat social information processing as a pathway through which biological hormones affect aggressive behavior. We examine the predictive effect of testosterone on aggressive behavior in juvenile offenders, the mediating role of hostile attention bias, and the moderating role of cortisol, thereby addressing how testosterone influences aggressive behavior and when this effect is more pronounced.

1.3.1 Testosterone and Aggressive Behavior: The Mediating Role of Attention Bias Biological hormone and social information processing perspectives each emphasize different direct effects on aggressive behavior, but the two perspectives are also connected, as biological hormones can indirectly influence aggressive behavior through social information processing.

Previous research has shown that exogenous testosterone increases can significantly enhance attention bias toward threatening faces (Terburg, Aarts, & Honk, 2012), which in turn leads individuals to perceive more hostility, make negative situational interpretations, and consequently exhibit aggressive behavior (Crick & Dodge, 1994). That is, attention bias is influenced by biological hormones while also affecting subsequent behavioral responses. Based on the direct effect of testosterone on aggressive behavior and the potential role of hostile attention bias, this study hypothesizes that testosterone can positively predict aggressive behavior in juvenile offenders (H1) and can also indirectly affect aggressive behavior through the mediating role of hostile attention bias (H2).

Notably, as an important step in cognitive processing, attention has significant dynamic characteristics, manifested as alternation and fluctuation between attention bias and attention avoidance (Zvielli, Bernstein, & Koster, 2015). However, traditional attention bias quantification methods based on the dot-probe paradigm (bias score, Dot-BS) can only reflect static, stable, and averaged attention bias characteristics at a specific time point after stimulus presentation (Kuckertz & Amir, 2015). Given the limitations of traditional bias score calculation and the dynamic nature of attention bias, Zvielli et al. (2015) proposed a trial-level bias score method (TL-BS) to reflect dynamic changes in attention bias. Research has found that TL-BS demonstrates higher split-half reliability and predictive power compared to traditional Dot-BS (Zvielli et al., 2015). Therefore, in addition to calculating traditional attention bias scores (Dot-BS), this study uses trial-level attention bias indices (TL-BS) to reflect the dynamic characteristics of attention bias and explore their associations with biological hormones and aggressive behavior.

1.3.2 The Moderating Role of Cortisol Although positive associations between testosterone and aggressive behavior and hostile attention bias have been supported by empirical research, findings have been inconsistent, with some studies reporting that testosterone is associated with lower competitiveness (Mehta & Josephs, 2010) and attentional avoidance of angry faces (van Peer, Enter, van Steenbergen, Spinhoven, & Roelofs, 2017). These inconsistencies suggest that potential moderating variables may influence these relationships. Researchers have begun to focus on the interaction between cortisol and testosterone, proposing the dual-hormone hypothesis (Mehta & Josephs, 2010).

The dual-hormone hypothesis suggests that as the final product of the HPA (hypothalamic-pituitary-adrenal) axis, cortisol can regulate the effects of testosterone by inhibiting testosterone secretion and blocking testosterone action pathways. Specifically, at low cortisol levels, testosterone shows a significant positive correlation with aggression, dominance, and control behaviors, whereas at high cortisol levels, this association is not significant (Mehta & Josephs, 2010). Empirical studies on male juvenile offenders (Popma et al., 2007), male adolescents (Grotzinger et al., 2018), and meta-analytic research (Dekkers et al., 2019) have all shown significant interactive effects between testosterone and cortisol in predicting social status, aggressive behavior, risk-taking, and dominance. Based on this, this study hypothesizes that cortisol moderates the direct effect of testosterone on aggressive behavior in juvenile offenders (H3).

It should be noted that although early research primarily focused on dual-hormone moderation of aggressive behavior, an increasing number of studies have shown that the interaction between cortisol and testosterone extends beyond this domain and may involve a range of cognitive and behavioral processes related to both hormones (Knight, Sarkar, Prasad, & Mehta, 2020). Research on decision-making (Prasad, Knight, & Mehta, 2019), episodic memory (Panizzon et al., 2018), and emotional control (Denson, Ronay, von Hippel, & Schira, 2013) have all provided new support for the dual-hormone hypothesis. Considering the inconsistent association between testosterone and hostile attention bias, the extensive interaction between the HPA and HPG axes, and the close relationships among testosterone, cortisol, and hostile attention bias, we propose that the dual-hormone hypothesis may also apply to hostile attention bias.

Specifically, cortisol may regulate the influence of testosterone on hostile attention bias by affecting testosterone secretion and action pathways. Therefore, this study hypothesizes that cortisol moderates the first half of the mediation pathway, influencing the effect of testosterone on hostile attention bias (H4).

1.4 The Current Study

In summary, based on the dual-hormone hypothesis and social information processing theory, this study integrates biological hormone and social information processing perspectives to investigate the influencing factors, mechanisms, and individual differences in violent aggressive behavior among juvenile offenders. This not only helps us understand the mechanisms through which biological hormones influence aggressive behavior but also provides implications for developing interventions against aggressive violence in juvenile offenders.

Specifically, this study proposes a moderated mediation model [Figure 1: see original paper] that examines both biological hormones and social information processing to investigate the mediating (hostile attention bias) and moderating (cortisol) mechanisms through which testosterone predicts aggressive behavior in juvenile offenders, addressing how testosterone influences aggressive behavior and when this effect is more pronounced. In terms of methodology, considering the dynamic characteristics of attention bias, this study calculates both traditional attention bias scores (Dot-BS) based on the dot-probe paradigm and trial-level bias scores (TL-BS) to reflect the dynamic attention bias process in juvenile offenders and compare the advantages of these two quantification methods.

[Figure 1: see original paper]

2. Method

2.1 Participants

Eighty-four male juvenile offenders were recruited from a juvenile correctional facility, with a mean age of 17.55 ± 0.52 years. Among all participants, 7.1% had sentences of one year, 73.2% had sentences of one to three years, and 19.7% had sentences of three to ten years. All participants had no history of psychiatric disorders or attention deficits, could read text correctly, had normal or corrected-to-normal vision, normal color vision, and normal intelligence.

This study was preregistered (https://osf.io/s4prd/). Before the study began, the experimenter informed participants about the tasks and compensation, after which all participants signed informed consent forms agreeing to participate.

2.2 Materials

2.2.1 Questionnaire Measures Chinese Version of the Buss-Perry Aggression Questionnaire (AQ-CV). The Buss-Perry Aggression Questionnaire is a self-report instrument measuring individual aggressive characteristics, comprising 29 items covering cognitive, emotional, and behavioral aspects of aggression. This study used the Chinese version revised by Li et al. (2011), which contains 30 items rated on a 5-point scale, including five subscales: physical aggression, verbal aggression, anger, hostility, and self-directed aggression. The internal consistency reliability of the total AQ-CV scale in this study was 0.91.

2.2.2 Biological Data Collection To control for circadian rhythms in salivary cortisol and testosterone secretion, all biological data were collected between 2:00 and 4:00 PM. Saliva samples (2 ml) were collected using salivary collection tubes to measure cortisol and testosterone levels. Samples were frozen at -20°C. Before analysis, samples were thawed and centrifuged at 3000 rpm for 10 minutes, and the supernatant was collected. Enzyme-Linked Immunosorbent Assay (ELISA) was used to determine salivary testosterone and cortisol concentrations. The minimum detection concentration for both testosterone and cortisol was less than 1.0 nmol/L (sensitivity), with intra- and inter-assay coefficients of variation less than 15%. All samples were assayed in duplicate. Normality tests indicated that both cortisol and testosterone levels were normally distributed.

2.2.3 Behavioral Experiments **Hostile Dot-Probe Task.** The dot-probe paradigm was used to measure attention bias toward hostile facial expressions. FaceGen 3.4 (http://FaceGen.com; Singular Inversions, 2009) was used to generate facial images as experimental materials. Forty college students rated 120 pictures of angry, disgusted, and fearful expressions on a 9-point hostility/aggression scale. The 10 male and 10 female faces with the highest hostility/aggression ratings were selected and paired with neutral expressions of the same faces, creating 20 hostile-neutral face pairs. Additionally, 20 neutral-neutral face pairs were constructed. These 40 face pairs served as experimental stimuli.

The experimental program was compiled using E-Prime 2.0 (Psychology Software Tools, Pittsburgh). The task consisted of a practice phase (10 trials) and a formal experiment (120 trials). After participants understood the instructions, the experiment began. A fixation point was presented at the center of the screen for 500 ms, followed by a face pair appearing on either side of the fixation point for 500 ms. Immediately after the faces disappeared, a probe letter "E" appeared at the location of one of the faces, and participants were required to respond as quickly and accurately as possible to the probe' s location. Participants with accuracy above 80% in the practice phase proceeded to the formal experiment; otherwise, they repeated the practice phase.

Hostile Aggression Stroop Task. An adapted emotional Stroop paradigm was used to investigate attention bias toward hostile aggression words. Forty college students rated 118 two-character hostile aggression words on a 9-point scale. The 40 words with the highest hostility/aggression ratings were selected.

Forty neutral words were selected from the *Modern Chinese Frequency Dictionary* to pair with the hostile words, forming the formal experimental materials. Additionally, five relatively hostile words and five neutral words were used as practice materials.

The experimental program was compiled using E-Prime 2.0, including a practice phase (10 trials) and a formal experiment (160 trials). After participants understood the instructions, the experiment began. A fixation point was presented at the center of the screen for 500 ms, followed by a stimulus (a red or green word) at the same location. Participants were instructed to ignore the word meaning and respond as quickly and accurately as possible to the word' s color. Practice phase accuracy above 80% was required to proceed to the formal experiment; otherwise, participants repeated the practice. Words used in practice were not used in the formal experiment.

2.3 Procedure

Participants first signed informed consent forms and then completed a questionnaire in the laboratory, including basic demographic and offense information and the Chinese version of the Buss-Perry Aggression Questionnaire. After questionnaire completion, saliva samples were collected for biological measurements. One day later, participants completed behavioral experiments including the hostile dot-probe task and hostile aggression Stroop task. The experimenter first introduced the procedure and instructions, demonstrated the operation method, and after participants were familiar with the rules, they completed the computer tasks. The entire process took approximately 30 minutes.

2.4 Data Processing

2.4.1 Data Cleaning For the hostile dot-probe task, data from one juvenile offender were missing, and five juvenile offenders had response accuracy below 80%; these participants' data were excluded. Additionally, practice phase data and all error responses (1.5%) were excluded. Correct response trials with reaction times below 200 ms or above 2000 ms (1.8%) and extreme data exceeding three standard deviations from each participant' s mean reaction time (1.8%) were also excluded. For the hostile aggression Stroop task, data from two juvenile offenders were missing. Practice phase data and all error responses (9.5%) were excluded, as were correct response trials with reaction times below 200 ms or above 2000 ms (0.5%) and extreme data exceeding three standard deviations from each participant' s mean reaction time (1.7%). The final sample included 76 valid participants.

2.4.2 Bias Score Calculation In the hostile aggression Stroop task, processing hostile words conflicts with processing font color. Longer reaction times to color indicate greater attentional and processing bias toward hostile words; shorter color-naming reaction times indicate attentional avoidance of hostile words (Domes et al., 2013). The bias score for the Stroop task (Stroop-BS) was calculated as the difference between participants' mean reaction times to hostile words and neutral words.

For the dot-probe task, bias scores were calculated using both the traditional static averaging method and the TL-BS method reflecting dynamic attention processes.

First, traditional attention bias scores (Dot-BS) were calculated to reflect static average characteristics, computed as participants' mean reaction time to incongruent trials (IT, where the probe location differed from the hostile expression location) minus mean reaction time to congruent trials (CT, where the probe location matched the hostile expression location). Dot-BS scores greater than zero indicate attentional bias toward hostile stimuli, while scores less than zero indicate attentional avoidance (Roy et al., 2008).

Second, the TL-BS method was used to reflect dynamic changes and overall processes in attention bias. Trial-level attention bias (TL-BS) was calculated by pairing congruent trials (CT) and incongruent trials (IT) within five trials of each other and computing the reaction time difference (IT-CT) for each pair (Zvielli et al., 2015). Compared with traditional attention bias indices, trial-level attention bias better captures the dynamic process of attention bias over time.

Five parameters can be further calculated from raw TL-BS scores: positive mean TL-BS (average of TL-BS values greater than zero), negative mean TL-BS (average of TL-BS values less than zero), positive peak TL-BS (maximum TL-BS value greater than zero), negative peak TL-BS (minimum TL-BS value less than zero), and TL-BS variability (sum of absolute TL-BS values divided by the number of TL-BS values). Positive mean TL-BS reflects average attentional bias toward hostile stimuli, negative mean TL-BS reflects average attentional avoidance of hostile stimuli, positive peak TL-BS reflects maximum attentional bias, negative peak TL-BS reflects maximum attentional avoidance, and TL-BS variability reflects attentional instability (manifested as alternating attentional bias and avoidance).

Using the 76 juvenile offenders as an example, traditional attention bias scores (Dot-BS) and trial-level attention bias scores (TL-BS) are shown in [Figure 2: see original paper]. Across all trials, the number of IT-CT pairs meeting the pairing criterion (within five trials of each other) was 18.41 (SD = 2.12, range: 12–20). As shown in the figure, compared with traditional attention bias scores, trial-level attention bias scores better reflect the attentional instability of juvenile offenders.

[Figure 2: see original paper]

2.5 Data Analysis

(1) Preliminary analyses were conducted, including descriptive statistics and correlation analyses; (2) Bias-corrected percentile bootstrap methods were

used to calculate confidence intervals, with 5,000 resamples and 95% confidence intervals. The PROCESS 2.16.3 macro for SPSS 20.0 (Hayes, 2012) was used to test the significance of simple mediation models, and based on significant simple mediation models, the significance of moderated mediation models was examined.

3. Results

3.1 Preliminary Analyses

Descriptive statistics and correlation analysis results are presented in . Testosterone showed significant positive correlations with cortisol and TL-BS variability, significant negative correlations with negative TL-BS parameters, and significant positive correlations with positive peak TL-BS. Regarding attention bias and aggression levels, negative mean TL-BS significantly predicted AQCV anger subscale scores; negative peak TL-BS significantly predicted AQCV total scores and physical aggression, verbal aggression, and anger subscale scores; variability significantly negatively predicted AQCV total scores, physical aggression subscale scores, and anger subscale scores.

3.2 Relationship Between Testosterone and Aggressive Behavior: Simple Mediation Model Test

First, Hayes' (2012) PROCESS macro was used to test the simple mediation model of attention bias. Based on correlation analysis results and controlling for age and education, models were constructed with attentional instability, negative peak TL-BS, and negative mean TL-BS as mediators, testosterone as the independent variable, and AQCV total scores, physical aggression subscale scores, verbal aggression subscale scores, and anger subscale scores as dependent variables. Multicollinearity tests indicated that all models had tolerance values greater than 0.67 and variance inflation factors less than 1.50, suggesting no serious multicollinearity issues.

Based on correlation analysis results, we hypothesized that attentional instability might mediate the relationship between testosterone and AQCV total scores, physical aggression subscale scores, and anger subscale scores. Using Model 4 in PROCESS 2.16.3 to test mediation effects, results are shown in . After including attentional instability in the model, testosterone significantly predicted attentional instability (β = 0.34, t = 3.24, p = 0.002), and attentional instability significantly predicted AQCV total scores, physical aggression subscale scores, and anger subscale scores (AQCV total: β = -0.29, t = -2.23, p = 0.03; physical aggression: β = -0.35, t = -2.67, p = 0.009; anger: β = -0.38, t = -2.89, p = 0.005). The bootstrap 95% confidence intervals for the indirect effects of attentional instability did not contain zero. Thus, attentional instability fully mediated the relationship between testosterone and AQCV total scores, physical aggression subscale scores, and anger subscale scores.

Based on correlation analysis results, we hypothesized that negative peak TL-BS might mediate the relationship between testosterone and AQCV total scores, physical aggression, verbal aggression, and anger subscale scores. With verbal aggression as the dependent variable, after including negative peak TL-BS in the model, testosterone marginally significantly predicted negative peak TL-BS (β = -0.22, t = -1.93, p = 0.058), and negative peak TL-BS significantly predicted verbal aggression subscale scores (β = 0.26, t = 2.10, p = 0.04). The bootstrap 95% confidence interval for the indirect effect of negative peak TL-BS did not contain zero (Effect = -0.03, SE = 0.01, 95% CI = -0.065 to -0.013). Thus, negative peak TL-BS fully mediated the relationship between testosterone and verbal aggression. However, the mediating effects of negative peak TL-BS in the prediction of AQCV total scores, physical aggression, and anger were not significant.

Based on correlation analysis results, we hypothesized that negative mean TL-BS would mediate the relationship between testosterone and anger subscale scores. Results showed that the bootstrap 95% confidence interval for the indirect effect of negative mean TL-BS contained zero, indicating a non-significant mediation effect.

3.3 Relationship Between Testosterone and Aggressive Behavior: Moderated Mediation Model Test

Based on simple mediation model results, with attentional instability as the mediator and cortisol as the moderator, we examined the predictive effect of testosterone on AQCV total scores, physical aggression subscale scores, and anger subscale scores. Results are shown in . The testosterone × cortisol interaction term significantly predicted only attentional instability (β = 0.30, t = 3.00, p = 0.004). Simple slope tests were used to further examine the specific interaction pattern, with results shown in Figure 3: see original paper. For participants with low cortisol levels, testosterone did not significantly predict attentional instability (simple slope = 0.34, t = 0.64, p = 0.53). For participants with high cortisol levels, testosterone significantly positively predicted attentional instability (simple slope = 2.47, t = 4.46, $p < 0.001$). Given that the direct effect of the independent variable on the dependent variable was not significant, Model 7 in PROCESS 2.16.3 was used to test the moderated mediation effect. Results showed that only at high cortisol levels did the simple mediation effect of attentional instability between testosterone and AQCV total scores, physical aggression subscale scores, and anger subscale scores become significant (see).

With negative peak TL-BS as the mediator and cortisol as the moderator, we examined the predictive effect of testosterone on verbal aggression subscale scores. Results are shown in . The testosterone × cortisol interaction term significantly predicted only negative peak TL-BS (β = -0.27, t = -2.51, p = 0.01). Simple slope tests were used to further examine the specific interaction pattern, with results shown in Figure 3: see original paper. For participants with low cortisol levels, testosterone did not significantly predict negative peak TL-BS (simple slope = -0.29, t = -0.10, p = 0.92). For participants with high cortisol levels, testosterone significantly negatively predicted negative peak TL-BS (simple slope = -9.95, t = -3.23, p = 0.002). Moderated mediation effect tests showed that only at high cortisol levels did the simple mediation effect of attentional instability between testosterone and verbal aggression subscale scores become significant (see).

[Figure 3: see original paper]

4. Discussion

This study integrated biological hormone and social information processing perspectives on aggressive behavior to examine the relationship between testosterone and aggressive behavior in juvenile offenders, the mediating role of hostile attention bias, and the moderating role of cortisol. Simple mediation analyses revealed that attentional instability fully mediated the relationship between testosterone and AQCV total scores, physical aggression subscale scores, and anger subscale scores, while negative peak TL-BS fully mediated the relationship between testosterone and verbal aggression subscale scores. Moderated mediation analyses showed that only at high cortisol levels were the mediating effects of attentional instability and negative peak TL-BS significant, such that testosterone could reduce aggression levels (AQCV total, physical aggression, and anger) by increasing attentional instability, and could also reduce verbal aggression by increasing attentional avoidance of hostile stimuli.

4.1 Testosterone and Aggressive Behavior

Contrary to our hypothesis, the study found no significant association between testosterone and aggressive behavior. Considering the nature of testosterone' s effect on aggressive behavior and the environmental context of juvenile correctional facilities, we propose that the special meaning of aggressive behavior in this environment may explain these results. As an endocrine hormone that motivates individuals to gain and maintain higher social status, the direction of testosterone' s behavioral effects depends on whether it facilitates social status improvement (Eisenegger et al., 2011; Terburg & van Honk, 2013), which is influenced by specific situational contexts (Carré & Archer, 2018). For example, in adolescent groups where aggressive behavior enhances social status, testosterone is positively associated with physical aggression (Tremblay, 1998). However, in socially interdependent contexts that value harmony and stability, testosterone is positively associated with prosocial behaviors such as honesty (Wibral, Dohmen, Klingmüller, Weber, & Falk, 2012) and reciprocity (Boksem et al., 2013). Thus, the association between testosterone and social behavior depends on the popularity rules in specific contexts (Dreher et al., 2016), which may help explain inconsistent findings regarding testosterone-aggression associations. For juvenile offenders in correctional facilities, aggressive behavior may no longer be an effective means to achieve higher social status; instead, showing lower aggression levels may be more advantageous. Moreover, 71.4% of participants in this study had served more than one year, suggesting that most individuals may have formed and adapted to new popularity norms. In this environment where aggressive behavior does not enhance social status, the positive direct effect of testosterone on aggressive behavior is weakened, resulting in no significant association between testosterone and aggression in this study.

4.2 The Mediating Role of Hostile Attention Bias

As mentioned above, in correctional facility environments, aggressive behavior may no longer be an effective way for individuals to improve social status, thus suppressing the relationship between testosterone and aggressive behavior. Building on this, mediation tests further revealed that testosterone could reduce aggression levels in juvenile offenders by positively predicting attentional avoidance and attentional instability toward hostile stimuli.

First, contrary to our hypothesis, we found that testosterone positively predicted attentional avoidance and attentional instability rather than attentional bias toward hostile stimuli. Considering testosterone' s role in promoting social status improvement and the importance of situational context in moderating testosterone-behavior relationships (Carré & Archer, 2018; Dreher et al., 2016), we propose that the special correctional facility environment may underlie these results. Testosterone' s effect on attention bias may be realized through amygdala activation. As a key brain structure for processing threatening stimuli, the amygdala may mediate testosterone' s influence on hostile attention bias. On one hand, the amygdala influences hostile attention bias characteristics (Buades-Rotger & Krämer, 2018). On the other hand, the amygdala' s androgen receptors make its response to angry faces susceptible to testosterone (Bos, Panksepp, Bluthé, & van Honk, 2012), and testosterone' s effect on the amygdala depends on individual motivation. Specifically, for individuals with approach motivation, increased testosterone enhances amygdala activation, strengthening attentional bias toward hostile stimuli. However, for individuals with avoidance motivation, testosterone inhibits amygdala activation, leading to more attentional avoidance (Radke et al., 2015). In correctional facility environments where aggressive behavior is not an effective means to improve social status, juvenile offenders may show more avoidance motivation, which causes testosterone to inhibit amygdala activation, resulting in greater attentional avoidance of hostile stimuli and a positive prediction of attentional avoidance by testosterone. The selective attention process is dynamic, manifested as alternation and fluctuation between attentional bias and avoidance (Zvielli et al., 2015), a characteristic more pronounced for special stimuli like hostile faces. Therefore, along with attentional avoidance, attentional instability also increases, manifested as positive prediction of attentional instability by testosterone.

Additionally, attentional avoidance and attentional instability toward hostile stimuli negatively predicted aggressive behavior, possibly reflecting emotion regulation processes and long-term dynamic characteristics in juvenile offenders. As an important emotion regulation strategy, attentional avoidance of hostile stimuli helps reduce sympathetic nervous system arousal (Gross, 2001), decreasing reactive aggression (Vasquez et al., 2013) and anger levels (Lievaart, Huijding, van der Veen, Hovens, & Franken, 2017). In this model, the negative predictive effect of attentional avoidance on aggressive behavior may reflect top-down emotion regulation and behavioral control processes. Building on this, the negative prediction of aggression by attentional instability represents dynamic changes in attention orientation during this process. Research shows that attentional regulation involves dynamic alternation between attentional bias and avoidance (Zvielli et al., 2015). Therefore, individuals showing high attentional instability can dynamically regulate their emotions, resulting in negative prediction of aggression by attentional instability.

4.3 The Moderating Role of Cortisol

Simple mediation models suggest that in correctional facility environments, the association between testosterone and aggressive behavior in juvenile offenders may be weakened or even reversed. Moderated mediation tests further revealed that cortisol moderated the relationship between testosterone and attention bias, with mediating effects of attentional avoidance and attentional instability significant only at high cortisol levels. In this process, cortisol may exert its moderating effect by increasing avoidance motivation toward hostile stimuli in juvenile offenders.

As a biological marker of punishment sensitivity, cortisol is involved in a series of stress-related metabolic and physiological reactions, causing individuals to experience more negative emotions and show more inhibitory and avoidance behaviors when presented with punishment stimuli (Smillie, Pickering, & Jackson, 2006). In correctional facilities, hostile, threatening, and aggression-related stimuli are often associated with punishment. When facing such stimuli, high cortisol levels result in higher avoidance motivation (Terburg, Morgan, & van Honk, 2009), which further causes testosterone to inhibit amygdala activation, ultimately manifesting as attentional avoidance of hostile stimuli (Radke et al., 2015). In summary, for juvenile offenders with high cortisol levels, their motivation to avoid negative stimuli combines with motivation to improve social status to cause testosterone to inhibit amygdala activation, showing attentional avoidance and consequently reducing aggressive behavior. For juvenile offenders with low cortisol levels, their avoidance motivation is relatively low, and only the motivation to improve social status weakens the association between testosterone and attention bias and aggressive behavior, ultimately resulting in non-significant mediation models.

4.4 Limitations and Clinical Implications

This study has several limitations: (1) Due to difficulties in obtaining data from juvenile offenders and quality issues with some participants' data, only 76 valid participants were included after screening, resulting in a relatively small sample size. (2) The current study only examined the effects of testosterone, cortisol, and hostile attention bias on aggressive behavior in juvenile offenders, while ignoring the potentially important roles of situational context and personal motivation. (3) The juvenile offenders in this study represent a unique population in terms of aggression levels and environmental background, making it impossible to design an appropriate control group (Clarke-McLean, 1996), and the current findings are limited to aggressive behavior of juvenile offenders in correctional facility environments.

The results also suggest that based on the mediating and moderating mechanisms of biological hormones influencing aggressive behavior, interventions for aggressive violence in juvenile offenders may target increasing attentional avoidance of hostile stimuli and elevating cortisol levels. First, given that biological hormones indirectly influence aggressive behavior through hostile attention bias, increasing attentional avoidance of hostile stimuli in juvenile offenders may effectively reduce aggression levels. Interventions for attentional processes have both "bottom-up" and "top-down" approaches. On one hand, "bottom-up" attentional processes can be targeted by strengthening individuals' implicit and automatic selective attention to non-threatening stimuli and adjusting their biased selection of external stimuli, such as through Attention Bias Modification Treatment (ABMT) (Hakamata et al., 2010). On the other hand, "top-down" attentional processes can be targeted by improving individuals' cognitive control and social information processing abilities to reduce attentional bias toward hostile stimuli, such as through cognitive behavioral therapy (Lee & DiGiuseppe, 2018). Second, considering the moderating role of cortisol, elevating cortisol levels may promote greater hostile attentional avoidance and consequently reduce aggression. Research has shown that family therapy effective in reducing aggression in children with antisocial behavior works by increasing cortisol levels (O' Neal et al., 2010).

5. Conclusion

This study yielded the following conclusions:

(1) Hostile attention bias (attentional instability, negative peak TL-BS) in juvenile offenders fully mediates the relationship between testosterone and aggressive behavior.

(2) Cortisol moderates the relationship between testosterone and hostile attention bias: only at high cortisol levels does testosterone positively predict attentional avoidance and attentional instability toward hostile stimuli, thereby reducing aggression levels.

Acknowledgments

We thank Tan Siqi, Shi Linlin, Chen Lin, Lin Huixian, and others for their assistance with data collection.

References

Archer, J. (2006). Testosterone and human aggression: an evaluation of the challenge hypothesis. *Neuroscience & Biobehavioral Reviews, 30*(3), 319-345.

Boksem, M. A., Mehta, P. H., Van den Bergh, B., van Son, V., Trautmann, S. T., Roelofs, K., …& Sanfey, A. G. (2013). Testosterone inhibits trust but promotes reciprocity. *Psychological Science, 24*(11), 2306-2314.

Bos, P. A., Panksepp, J., Bluthé, R. M., & van Honk, J. (2012). Acute effects of steroid hormones and neuropeptides on human social-emotional behavior: a review of single administration studies. *Frontiers in Neuroendocrinology, 33*(1), 17-35.

Brooks, J. H., & Reddon, J. R. (1996). Serum testosterone in violent and nonviolent young offenders. *Journal of Clinical Psychology, 52*(4), 475-483.

Buades-Rotger, M., & Krämer, U. M. (2018). From words to action: implicit attention to antisocial semantic cues predicts aggression and amygdala reactivity to angry faces in healthy young women. *Aggressive Behavior, 44*(6), 624-637.

Carré, J. M., & Archer, J. (2018). Testosterone and human behavior: the role of individual and contextual variables. *Current Opinion in Psychology, 19*, 149-153.

Clarke-McLean, J. G. (1996). Social networks among incarcerated juvenile offenders. *Social Development, 5*(2), 203-217.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children' s social adjustment. *Psychological Bulletin, 115*(1), 74-101.

Dabbs J. M., Carr, T. S., Frady, R. L., & Riad, J. K. (1995). Testosterone, crime, and misbehavior among 692 male prison inmates. *Personality and Individual Differences, 18*(5), 627-633.

Dekkers, T. J., van Rentergem, J. A. A., Meijer, B., Popma, A., Wagemaker, E., & Huizenga, H. M. (2019). A meta-analytical evaluation of the dual-hormone hypothesis: does cortisol moderate the relationship between testosterone and status, dominance, risk taking, aggression, and psychopathy? *Neuroscience & Biobehavioral Reviews, 96*, 250-271.

Denson, T. F., Ronay, R., von Hippel, W., & Schira, M. M. (2013). Endogenous testosterone and cortisol modulate neural responses during induced anger control. *Social Neuroscience, 8*(2), 165-177.

Domes, G., Mense, J., Vohs, K., & Habermeyer, E. (2013). Offenders with antisocial personality disorder show attentional bias for violence-related stimuli. *Psychiatry Research, 209*(1), 78-82.

Dreher, J. C., Dunne, S., Pazderska, A., Frodl, T., Nolan, J. J., & O' Doherty, J. P. (2016). Testosterone causes both prosocial and antisocial status-enhancing behaviors in human males. *Proceedings of the National Academy of Sciences, 113*(41), 11633-11638.

Eisenegger, C., Haushofer, J., & Fehr, E. (2011). The role of testosterone in social interaction. *Trends in Cognitive Sciences, 15*(6), 263-271.

Fernàndez-Castillo, N., & Cormand, B. (2016). Aggressive behavior in humans: genes and pathways identified through association studies. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 171*(5), 676-696.

Gross, J. J. (2001). Emotion regulation in adulthood: timing is everything. *Current Directions in Psychological Science, 10*(6), 214-219.

Grotzinger, A. D., Mann, F. D., Patterson, M. W., Tackett, J. L., Tucker-Drob, E. M., & Harden, K. P. (2018). Hair and salivary testosterone, hair cortisol, and externalizing behaviors in adolescents. *Psychological Science, 29*(5), 688-699.

Hakamata, Y., Lissek, S., Bar-Haim, Y., Britton, J. C., Fox, N. A., Leibenluft, E., …& Pine, D. S. (2010). Attention bias modification treatment: a meta-analysis toward the establishment of novel treatment for anxiety. *Biological Psychiatry, 68*(11), 982-990.

Hayes, A. F. (2012). *PROCESS: a versatile computational tool for observed variable mediation, moderation, and conditional process modeling* [White paper]. Retrieved from http://www.afhayes.com/public/process2012.pdf

Huesmann, L. R. (2018). An integrative theoretical understanding of aggression: a brief exposition. *Current Opinion in Psychology, 19*, 119-124.

Knight, E. L., Sarkar, A., Prasad, S., & Mehta, P. H. (2020). Beyond the challenge hypothesis: the emergence of the dual-hormone hypothesis and recommendations for future research. *Hormones and Behavior, 104657*.

Kuckertz, J. M., & Amir, N. (2015). Attention bias modification for anxiety and phobias: current status and future directions. *Current Psychiatry Reports, 17*(2), 9.

Laue, C., Griffey, M., Lin, P. I., Wallace, K., Van der Schoot, M., Horn, P., … & Barzman, D. (2018). Eye gaze patterns associated with aggressive tendencies in adolescence. *Psychiatric Quarterly, 89*(3), 747-756.

Lee, A. H., & DiGiuseppe, R. (2018). Anger and aggression treatments: a review of meta-analyses. *Current Opinion in Psychology, 19*, 65-74.

Li, X. Y., Fei, L. P., Zhang, Y. L., Niu, Y. J., Tong, Y. S., & Yang, S. J. (2011). Development, reliability and validity of the Chinese version of Buss & Perry Aggression Questionnaire. *Chinese Journal of Nervous and Mental Diseases, 37*(10), 607-613.

Lievaart, M., Huijding, J., van der Veen, F. M., Hovens, J. E., & Franken, I. H. (2017). The impact of angry rumination on anger-primed cognitive control.

Journal of Behavior Therapy and Experimental Psychiatry, 54, 135–142.

Lu, Q., Guo, K. Y., Zhang, M., Zhang, X. B., Hu, F. Q., & Yang, J. L. (2018). Chinese juvenile delinquency research report 2017: based on a comparative study of juvenile offenders and other groups. *Juvenile Delinquency, 219*(6), 31–45.

Manning, K. E. (2019). Seeing red? A systematic review of the evidence for attentional biases to threat-relevant stimuli in propensity to reactive aggression. *Aggression and Violent Behavior, 47*, 90–103.

Mehta, P. H., & Josephs, R. A. (2010). Testosterone and cortisol jointly regulate dominance: evidence for a dual-hormone hypothesis. *Hormones and Behavior, 58*(5), 898–906.

Miller, N. V., & Johnston, C. (2019). Social threat attentional bias in childhood: relations to aggression and hostile intent attributions. *Aggressive Behavior, 45*(3), 245–254.

Montoya, E. R., Terburg, D., Bos, P. A., & Van Honk, J. (2012). Testosterone, cortisol, and serotonin as key regulators of social aggression: a review and theoretical perspective. *Motivation and Emotion, 36*(1), 65–73.

O' Neal, C. R., Brotman, L. M., Huang, K. Y., Gouley, K. K., Kamboukos, D., Calzada, E. J., & Pine, D. S. (2010). Understanding relations among early family environment, cortisol response, and child aggression via a prevention experiment. *Child Development, 81*(1), 290–305.

Panagiotidis, D., Clemens, B., Habel, U., Schneider, F., Schneider, I., Wagels, L., & Votinov, M. (2017). Exogenous testosterone in a non-social provocation paradigm potentiates anger but not behavioral aggression. *European Neuropsychopharmacology, 27*(11), 1172–1184.

Panizzon, M. S., Hauger, R. L., Xian, H., Jacobson, K., Lyons, M. J., Franz, C. E., & Kremen, W. S. (2018). Interactive effects of testosterone and cortisol on hippocampal volume and episodic memory in middle-aged men. *Psychoneuroendocrinology, 91*, 115–122.

Popma, A., Vermeiren, R., Geluk, C. A., Rinne, T., van den Brink, W., Knol, D. L., ···& Doreleijers, T. A. (2007). Cortisol moderates the relationship between testosterone and aggression in delinquent male adolescents. *Biological Psychiatry, 61*(3), 405–411.

Prasad, S., Knight, E. L., & Mehta, P. H. (2019). Basal testosterone' s relationship with dictator game decision-making depends on cortisol reactivity to acute stress: a dual-hormone perspective on dominant behavior during resource allocation. *Psychoneuroendocrinology, 101*, 150–159.

Radke, S., Volman, I., Mehta, P., van Son, V., Enter, D., Sanfey, A., ···& Roelofs, K. (2015). Testosterone biases the amygdala toward social threat approach. *Science Advances, 1*(5), e1400074.

Rosell, D. R., & Siever, L. J. (2015). The neurobiology of aggression and violence. *CNS Spectrums, 20*(3), 254–279.

Roy, A. K., Vasa, R. A., Bruck, M., Mogg, K., Bradley, B. P., Sweeney, M., ···& CAMS Team. (2008). Attention bias toward threat in pediatric anxiety disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(10), 1189–1196.

Smillie, L. D., Pickering, A. D., & Jackson, C. J. (2006). The new reinforcement sensitivity theory: implications for personality measurement. *Personality and Social Psychology Review, 10*(4), 320–335.

Terburg, D., & van Honk, J. (2013). Approach–avoidance versus dominance-submissiveness: a multilevel neural framework on how testosterone promotes social status. *Emotion Review, 5*(3), 296–302.

Terburg, D., Aarts, H., & van Honk, J. (2012). Testosterone affects gaze aversion from angry faces outside of conscious awareness. *Psychological Science, 23*(5), 459–463.

Terburg, D., Morgan, B., & van Honk, J. (2009). The testosterone–cortisol ratio: a hormonal marker for proneness to social aggression. *International Journal of Law and Psychiatry, 32*(4), 216–223.

Tremblay, R. E. (1998). Testosterone, physical aggression, dominance, and physical development in early adolescence. *International Journal of Behavioral Development, 22*(4), 753–777.

van Peer, J. M., Enter, D., van Steenbergen, H., Spinhoven, P., & Roelofs, K. (2017). Exogenous testosterone affects early threat processing in socially anxious and healthy women. *Biological Psychology, 129*, 82–89.

Vasquez, E. A., Pedersen, W. C., Bushman, B. J., Kelley, N. J., Demeestere, P., & Miller, N. (2013). Lashing out after stewing over public insults: the effects of public provocation, provocation intensity, and rumination on triggered displaced aggression. *Aggressive Behavior, 39*(1), 13–29.

Wibral, M., Dohmen, T., Klingmüller, D., Weber, B., & Falk, A. (2012). Testosterone administration reduces lying in men. *PloS One, 7*(10).

Wilkowski, B. M., & Robinson, M. D. (2010). The anatomy of anger: an integrative cognitive model of trait anger and reactive aggression. *Journal of Personality, 78*(1), 9–38.

Yang, B., Xiao, Y. Q., Lai, L., & Wu, M. X. (2018). Evidence-based modification of juvenile violent delinquency. *Juvenile Delinquency Prevention Research, 1*, 32–42.

Zvielli, A., Bernstein, A., & Koster, E. H. (2015). Temporal dynamics of attentional bias. *Clinical Psychological Science, 3*(5), 772–788.

Note: Figure translations are in progress. See original paper for figures.