Overconfidence in wargames: experimental evidence on expectations, aggression, gender and testosterone

Dominic D. P. Johnson[1,*,†]**, Rose McDermott**[2]**, Emily S. Barrett**[3]**,
Jonathan Cowden**[2]**, Richard Wrangham**[3]**, Matthew H. McIntyre**[4]
and Stephen Peter Rosen[5]

[1]*Society of Fellows and Woodrow Wilson School of Public and International Affairs, 222 Bendheim Hall,
Princeton University, Princeton, NJ 08544, USA*
[2]*Department of Political Science, University of California Santa Barbara, Santa Barbara, CA 93106, USA*
[3]*Department of Anthropology, Harvard University, Peabody Museum, 11 Divinity Avenue,
Cambridge, MA 02138, USA*
[4]*Department of Epidemiology, Harvard University, 677 Huntington Avenue, Boston, MA 02115, USA*
[5]*Olin Institute for Strategic Studies, Harvard University, 1033 Massachusetts Avenue, Cambridge, MA 02138, USA*

Overconfidence has long been noted by historians and political scientists as a major cause of war. However, the origins of such overconfidence, and sources of variation, remain poorly understood. Mounting empirical studies now show that mentally healthy people tend to exhibit psychological biases that encourage optimism, collectively known as 'positive illusions'. Positive illusions are thought to have been adaptive in our evolutionary past because they served to cope with adversity, harden resolve, or bluff opponents. Today, however, positive illusions may contribute to costly conflicts and wars. Testosterone has been proposed as a proximate mediator of positive illusions, given its role in promoting dominance and challenge behaviour, particularly in men. To date, no studies have attempted to link overconfidence, decisions about war, gender, and testosterone. Here we report that, in experimental wargames: (i) people are overconfident about their expectations of success; (ii) those who are more overconfident are more likely to attack; (iii) overconfidence and attacks are more pronounced among males than females; and (iv) testosterone is related to expectations of success, but not within gender, so its influence on overconfidence cannot be distinguished from any other gender specific factor. Overall, these results constitute the first empirical support of recent theoretical work linking overconfidence and war.

Keywords: overconfidence; aggression; gender; testosterone; war; positive illusions

1. INTRODUCTION

Decision-makers sometimes choose war even when they do not expect to win. They may, for example, anticipate third party intervention or an improvement in their bargaining position (Paul 1994). But even these scenarios do not always solve a central problem in international relations dubbed the 'war puzzle': rational states—whether expectant winners or expectant losers—should not fight because if they assess each other accurately, they could avoid the costs and risks of war (blood, treasure and uncertainty) by negotiating a pre-war bargain reflecting their relative power (Fearon 1995). Because wars do occur, states appear to overestimate their relative power. Indeed, a recurrent theme among studies of the causes of war is that overconfidence is frequently associated with the outbreak of violence (Blainey 1973; Jervis 1976; Lebow 1981; Stoessinger 1998; Van Evera 1999; Ganguly 2001; Johnson 2004). However, the origin of this overconfidence is not understood (political scientists have tended to seek answers in the shortcomings of information, bureaucratic processes or institutional biases).

A potential solution to the war puzzle derives from an empirical feature of human nature: most normal people exhibit cognitive and motivated biases towards: (i) self-aggrandizement; (ii) an illusion of control over events; and (iii) invulnerability to risk—three widely replicated and robust phenomena collectively known as 'positive illusions' (Taylor & Brown 1988, 1994; Peterson 2000). There are reasons to believe that such individual biases are likely to be further exacerbated at group, organizational and societal levels (Janis 1972; LeShan 2002; Van Evera 2003; Johnson & Tierney 2006). Although positive illusions amount to systematic errors in assessment, they confer numerous advantages in many life tasks as a kind of self-fulfilling prophecy—promoting health, creativity, physical and mental performance in the face of otherwise debilitating obstacles (Taylor 1989; Taylor & Armor 1996; Gillham 2000). Various authors have argued that such adaptive advantages led to a selection pressure for positive

illusions through our evolutionary history (Tiger 1979; Trivers 2000; Nettle 2004; Haselton & Nettle 2006). One proposed adaptive advantage is specifically linked to conflict: positive illusions may have improved combat performance in the past by bolstering resolve and/or deceiving opponents via bluffing (Wrangham 1999; Johnson *et al.* 2002; Johnson 2004). This would predict that positive illusions are greater under threat, and are stronger in males (who have been the predominant warriors and fighters throughout evolutionary history).

While positive illusions may have been adaptive in our environment of evolutionary adaptation, present day stimuli and feedback that are evolutionarily novel may sometimes allow them to wreak havoc. A number of lines of evidence corroborate the stereotype that men (particularly young men), and not women, are susceptible to unwarranted levels of perceived invulnerability and confidence in their ability, and testosterone has been proposed as a candidate gender-biased proximate mechanism; in situations of conflict testosterone levels tend to rise, and this increases the probability of confrontational behaviour which may lead to violence (Wrangham & Peterson 1996; Baumeister & Boden 1998; Mazur & Booth 1998; Muller & Wrangham 2001; Rosen 2004). In positions of political and military power—which are held predominantly by men—overconfidence may lead to less compromise, more conflict, and more costly and/or more frequent wars (Johnson 2004). Here, we close the gap between these theoretical propositions and the real world using data from a wargame specifically designed to analyse decisions within an international conflict scenario. We were interested in whether, when, and which players made 'unprovoked attacks' during the game, where unprovoked attacks were defined as launching a war without any prior violence carried out by the other side. We test four hypotheses:

(i) people are overconfident about their expectations of success in conflict;
(ii) those who are more overconfident are more likely to make unprovoked attacks;
(iii) overconfidence and unprovoked attacks are more pronounced among males; and
(iv) overconfidence and unprovoked attacks are correlated with testosterone.

2. MATERIAL AND METHODS

Following an existing wargame methodology (McDermott & Cowden 2001, submitted), 200 experimental subjects played a simulated crisis game in a networked computer laboratory. Each person was asked to role-play the leader of a fictitious country in conflict with another over newly discovered diamond resources along a disputed border. Subjects were paid $20 to participate, and an additional $10 if they won the game (defined as finishing with the greatest industrial wealth, or being the sole surviving state if they defeated their opponent in war). Subjects were given background information on the scenario and were asked to resolve the crisis without being told how to do so. Each player played the entire game via their own private computer terminal. No dominant strategy for success existed; subjects could, and did, win through a variety of mechanisms from negotiation to war. All subjects were randomly assigned to dyads (pairs of two that

would play against each other). They were not aware of the identity or sex of their opponent. Players were not able to size up their opponents before play. There were 20–40 people at any one time in the room, and because of random pairing, it was impossible for anyone to know which of the others they were playing. Also, subjects were playing via the web and were told that they may or may not be playing someone in the room. Similarly, subjects were not aware of how long the game would last. In actuality, subjects always interacted over the course of six rounds. Each round lasted 7 min, and participants were made aware of these time limits by clocks which ran in the lower right-hand corner of their computer screens.

At the start and again in each subsequent round, subjects were given 100 million dollars which they could allocate as they chose. They could invest in their military, in industrial infrastructure, or they could reserve the money in cash. Subjects remained aware from the outset that the player with the most money in their 'industrial account' at the end of the game would emerge victorious. If they negotiated successfully, they could gain additional resources from the sale of the diamonds; they could also gain resources and, thereby, wealth by achieving success in battle. Victory in battle was determined by the computer program according to prior probabilities, which the subjects knew in advance from the background material they read at the start of the game. Victory was dependent, in part, on how many battalions the player committed to battle, but there was also an element of chance (akin to rolling a die). Some dyads terminated prior to six rounds of play because they destroyed each other in combat.

Each round of play, subjects had to make a decision about what action to take. These actions included doing nothing, negotiating, making a threat, initiating or continuing war, or surrendering. In addition, each person sent a written message to the adversary. At the end of each round, these messages were displayed to the opponent.

Before the game, players forecast their own rank—how they expected to perform in the game compared to all 200 subjects playing the game (hereafter 'pre-game self-ranking'). *After* having played the game, but before the results were known, players again forecast their own rank—how they thought they had performed in the game compared to all 200 subjects playing the game (hereafter 'post-game self-ranking'). Although the games were always played in pairs of two subjects, every individual's score in the game was ranked against the whole sample of subjects (similarly, players were asked to rank their expected personal performance among the whole sample of subjects). Each players' actual rank following the game was determined according to their final industrial production score, with ties allowed.

Players had the option of launching wars at any point. Unprovoked attacks were recorded wherever a player launched a war without any prior violence being carried out by the other side. Because decisions were revealed simultaneously in each round, both could decide to attack within the same round of the game, in which case both players would be recorded as having made an unprovoked attack. Retaliatory attacks in subsequent rounds, after having been attacked oneself, did not count as unprovoked attacks.

(a) Subject groups

There were 186 subjects returning complete data in this analysis, 107 men and 79 women. They came from an existing subject pool administered by the Harvard Business School

experimental laboratory. These subjects were drawn from the larger Cambridge, MA area and included (but not exclusively) undergraduates, graduate students, and staff from a number of colleges and universities in the area. The average age was 22.3 years (median 21.0, range 18–65, s.d. = 5.5). Our sample was 61% Caucasian, 20% Asian or Asian–American, 11% African American, 3% Hispanic, 1% Native American and 4% other. Among all subjects, 69% were single, 22% dating, 3% living with a partner and 5% married. Comparing education levels, 1% had completed high school only, 74% had some college education, 6% were college graduates, 13% had some graduate school education, and 7% had completed graduate school. All subjects gave informed consent for this experiment, and we followed Institutional Review Board experimental protocol for human subject research.

(b) *Testosterone assays*
Our experiment was designed to limit variation in testosterone over: life-span (age was recorded and controlled for in analyses involving testosterone); the circadian rhythm (all experiments were run in the afternoon); and circannual rhythms (all experiments were conducted in a single 3 day period in the spring of 2003). Subjects gave three saliva samples during the course of the study: immediately upon entering the computer laboratory, after three rounds of play, and at the conclusion of play. In instances that the game ended before three rounds of play, players were asked to give a final saliva sample immediately after play stopped, for a total of two, rather than three samples. Saliva collection procedures followed previously validated methods (Lipson & Ellison 1989; Granger *et al.* 1999). Subjects were given a stick of Extra Original Flavour gum to stimulate saliva flow and were then asked to salivate into a 15 ml collection vial which had been pre-treated with sodium azide, an anti-bacterial agent. The samples were temporarily stored at room temperature, after which they were frozen, then thawed 24 hours prior to being assayed.

Samples were assayed for testosterone at the Harvard University Reproductive Ecology Laboratory, following a modified version of a ^{125}I-based, double-antibody radio-immunoassay kit (DSL-4100) produced by Diagnostic Systems Laboratories, Inc. (Webster, TX). This adapted protocol is described in greater detail elsewhere (Burnham *et al.* 2003; Gray *et al.* 2004).

Subject samples were assigned to eight different assay groups; assays 1–3 were exactly gender balanced, whereas 4–8 had slightly more males than females (with no reason to bias results). Efforts were made to measure the testosterone levels of both partners in a dyad within the same assay to make their results more directly comparable. The interassay coefficients of variation for the standard reference low-T and high-T pools were 20.2 and 5.3%, respectively.

3. RESULTS
Of all 1080 decisions made by all players during the six rounds of the wargames, 70.7% were to negotiate (note that this could be after a period of warfare), 19.6% were to do nothing, 5.9% to fight, 3.7% to make a threat (via a written message) and 0% surrenders. Wars occurred in 47.8% of the games (i.e. situations in which one or the other side, or both, attacked the other at some point during the six rounds). The mean testosterone level among females was 74.9 pmol l^{-1} (picomoles per litre; s.d. = 53.4), and among males was 341.5 pmol l^{-1}
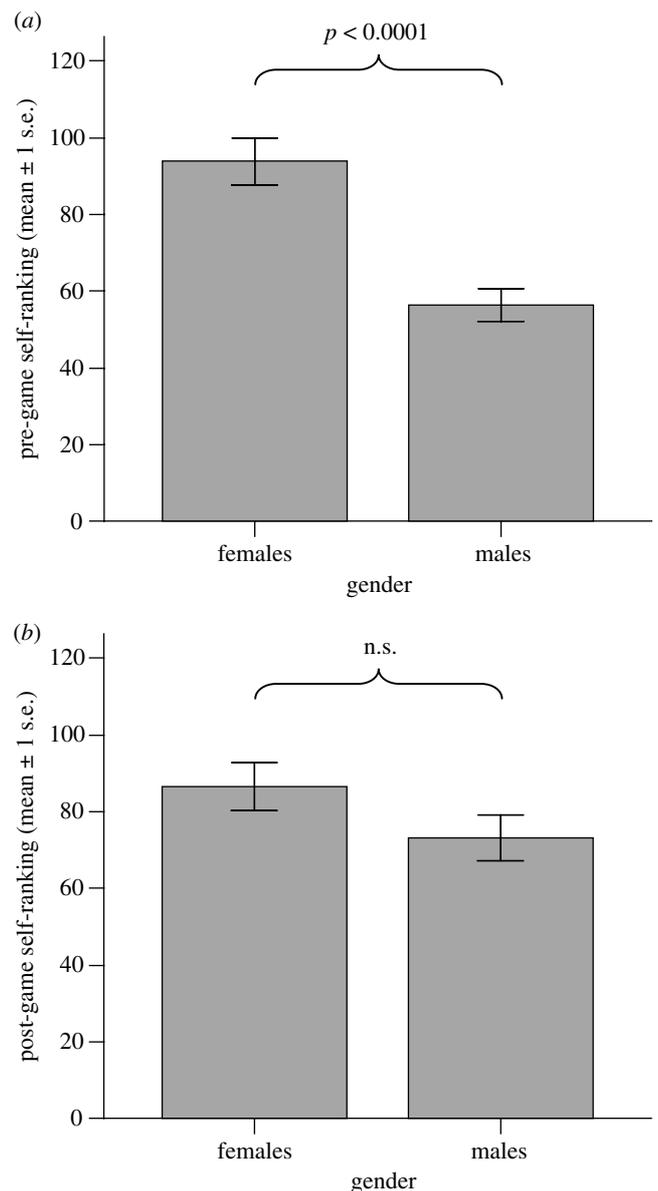

Figure 1. Self-rankings of expected success for both genders (mean ranks ± 1 s.e.), reported either: (a) pre-game; or (b) post-game (but before players knew the results). Note that a lower number corresponds to a higher expected rank. Self-rankings were significantly lower than 100 ($p < 0.05$) in all cases except for females' pre-game self-ranking.

(s.d. = 202.1). We used non-parametric statistics throughout because data were not normally distributed, and because correlations were not necessarily expected to be linear. The exception was in our analyses of testosterone using linear multiple regression, when the relevant data were normalized via a transformation to their natural logarithm.

(a) *Hypothesis 1*
Were people overconfident about their expectations of success? Prior to playing the game, the mean pre-game self-ranking was 72.3 (where 1 represents expecting to be the best of all players, and 200 the worst). Players' expectations of their performance were significantly above the middle value of 100 (Wilcoxon signed-rank test: $Z = -6.36$, $p < 0.0001$; the middle value is 100 because although only 186 subjects returned complete data, players were ranking themselves out of a publicly stated 200 competitors at the
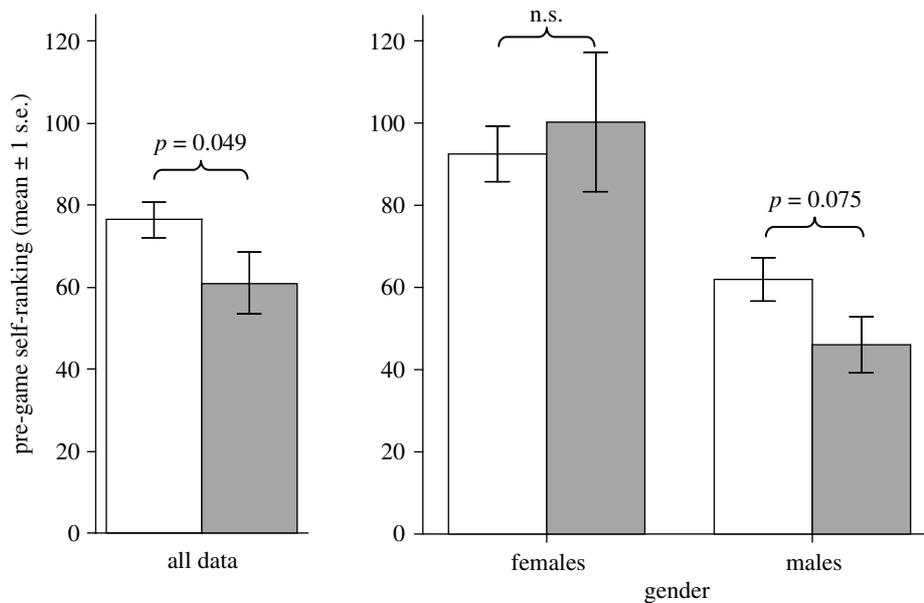
Figure 2. Pre-game self-rankings and unprovoked attacks, for all data and when split by gender (mean ranks ± 1 s.e.). Grey bars represent players who launched unprovoked attacks, white bars represent players who did not.

start). Figure 1 shows this effect was largely due to males (see Hypothesis 3 below for details of gender differences).

Our results cannot be accounted for by certain individuals correctly assessing their superiority, because pre-game self-rankings did not correlate positively with actual rank (in fact, there was a significant *negative* correlation: Spearman's $\rho = -0.16$, $p = 0.04$), nor with actual score ($\rho = 0.01$, $p = 0.91$). These results were similar when split by gender (using actual rank, the correlation approached significance only among males, although it was negative in both sexes: males, $\rho = -0.18$, $p = 0.08$; females, $\rho = -0.10$, $p = 0.39$; using actual score, for both sexes $\rho < 0.08$, and $p > 0.44$). If anything, therefore, those who expected to do best, tended to do worst. This suggests that positive illusions were not only misguided, but actually may have been detrimental to performance in this scenario.

(b) Hypothesis 2

Were those who were more overconfident more likely to make unprovoked attacks? Those who carried out unprovoked attacks on their opponents gave significantly higher pre-game self-rankings (Mann–Whitney U-test: $Z = 1.97$, $N = 47,137$, $p = 0.049$). Split by gender, this effect was no longer statistically significant, although a positive trend among men was still apparent (males: $Z = 1.78$, $N = 34,72$, $p = 0.075$; females: $Z = 0.46$, $N = 13,65$, $p = 0.64$; see figure 2). There was no relationship between pre-game self-ranking and the propensity to retaliate *after* having been attacked (both sexes: $Z = 0.24$, $N = 8,7$, $p = 0.81$; note the small sample size of this outcome, which meant there were too few instances of retaliation for a reliable test when split by gender).

(c) Hypothesis 3

Were overconfidence and unprovoked attacks more pronounced among males than females? The overconfidence examined in Hypothesis 1 was, in fact, solely due to males. Female pre-game self-rankings did not differ significantly from 100 (Wilcoxon signed-rank test:

$Z = -1.07$, $p = 0.28$; mean $= 93.9$). However, males' pre-game self-rankings were significantly above 100 (Wilcoxon signed-rank test: $Z = -7.15$, $p < 0.0001$; mean $= 56.5$). As well as this 'above average' effect among males, the male–female difference in self-ranking was also significant (Mann–Whitney U-test: $Z = 4.73$, $N = 107,79$, $p < 0.0001$; see figure 1a). After the experiment (that is, after having played the game but before knowing their own or the overall distribution of results), male self-rankings were still significantly greater than 100, though to a lesser extent (Wilcoxon signed-rank test: $Z = -3.90$, $p < 0.0001$; mean $= 73.2$). At this post-game stage, female self-rankings were also slightly but significantly greater than 100 (Wilcoxon signed-rank test: $Z = -1.99$, $p = 0.047$; mean $= 86.5$), and as a result the gender difference was no longer statistically significant (Mann–Whitney U-test: $Z = 1.59$, $N = 99,76$, $p = 0.11$; see figure 1b). This suggests that, in the process of the wargame, males learned that they were not as good as they initially forecast, and females learned that they were better than they forecast. But both genders now exhibited a degree of overconfidence in terms of ranking their performance above average.

Males also made significantly more unprovoked attacks than females ($\chi^2 = 5.61$, d.f. $= 1$, $p = 0.018$), and were significantly more likely to either attack or retaliate rather than never fight ($\chi^2 = 4.24$, d.f. $= 1$, $p = 0.040$). Even though players were unaware of the identity or gender of their opponent, wars and unprovoked attacks were highest among male–male dyads, next most common among mixed dyads, and least common in female–female dyads.

(d) Hypothesis 4

Were overconfidence and unprovoked attacks correlated with testosterone? Testosterone levels sampled at the outset were significantly related to pre-game self-rankings (Spearman's $\rho = -0.33$, $N = 180$, $p < 0.0001$; later testosterone samples were *not* related to *post*-game self-rankings, whether all data or split by gender: all $\rho < 0.13$, all $p > 0.11$). However, the effect disappeared when examining males and females as separate groups (males,
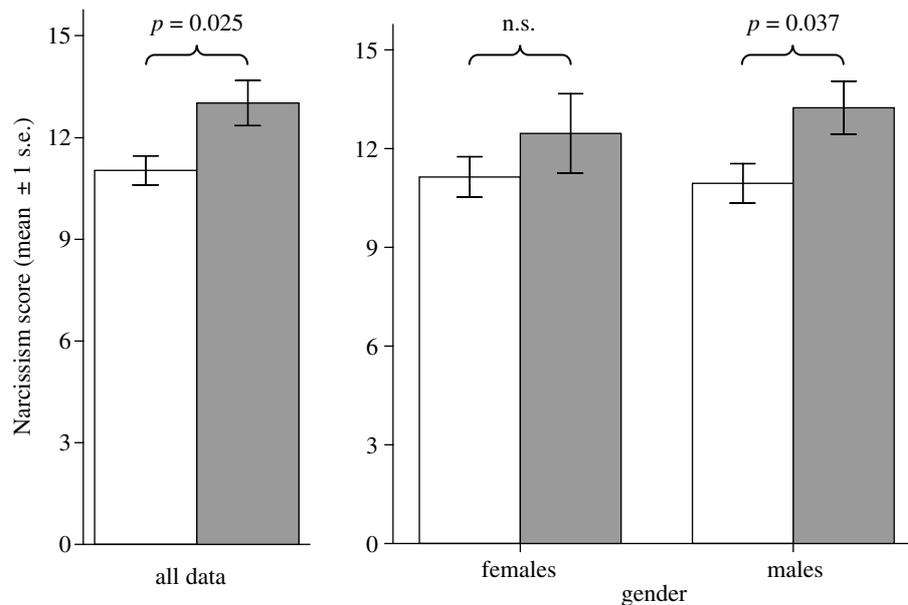
Figure 3. Narcissism score and unprovoked attacks, for all data and when split by gender (mean score ± 1 s.e.). Grey bars represent players who launched unprovoked attacks, white bars represent players who did not.

$\rho = 0.03$, $N = 102$, $p = 0.79$; females, $\rho = -0.16$, $N = 78$, $p = 0.15$). A multiple linear regression, removing variation due to gender, confirmed that testosterone did not account for any significant remaining variation in pre-game self-ranking (partial correlation coefficient: $t = -0.99$, $p = 0.32$; nor when examining testosterone levels sampled at the mid-point or at the end of the game against *post*-game self-rankings). For these tests, testosterone data were normalized with a transformation to the natural logarithm; self-ranking data were nearly normal except that a minority of subjects expected to gain very high ranks, creating a bimodal distribution that was not possible to improve. Age made no difference in any of these tests. Age data were heavily skewed with a mean of 22.3 and range 18–65 (s.d. $= 5.5$). We excluded four outlier subjects with ages greater than three standard deviations from the mean (that is, 39 years or older), and normalized age to its natural logarithm (this transformation could not be improved upon). As expected, there was a negative correlation between ln(age) and pre-game ln(testosterone) (Pearson's $r = -0.17$, $p = 0.027$). However, this did not affect the conclusions above: controlling for age but not gender, testosterone remained a significant predictor of pre-game self-ranking ($F(2,173) = 14.28$, $p < 0.0001$; partial correlation coefficient for ln(testosterone): $t = -5.03$, $p < 0.0001$). When gender was included, the overall model was still significant ($F(3,172) = 14.32$, $p < 0.0001$). However, in this model, variation in pre-game self-ranking could be attributed to gender (partial: $t = -3.54$, $p < 0.001$), and ln(age) (partial $t = -3.15$, $p < 0.002$), but *not* to testosterone (partial: $t = -0.65$, $p = 0.52$). Actual rank or final score did not correlate with raw testosterone levels either (both $\rho < 0.08$, both $p > 0.32$), nor with ln(testosterone) when controlling for age and gender. Thus, although in the pooled data testosterone is a significant correlate of pre-game self-ranking, there is no evidence that testosterone has an independent effect on expectations of success over and above the effect of gender. There was also no evidence that those with higher testosterone were more likely to make unprovoked attacks

(Mann–Whitney U-test: $Z = 1.64$, $N = 45,133$, $p = 0.10$), nor when males and females were examined separately (both $Z < 0.60$, both $p > 0.55$), nor when using testosterone sampled at the mid-point or at the end of the game (for all data and when split by gender: all $Z < 1.77$, all $p > 0.07$).

Finally, in probing the characteristics of individuals that were prone to overconfidence and launching wars, we found that levels of narcissism (as measured by the Narcissistic Personality Inventory, Raskin & Terry 1988) were significantly related to pre-game self-rankings. Males (but not females) with high narcissistic qualities tended to expect to do better (all data, Spearman's $\rho = -0.21$, $N = 185$, $p = 0.005$; males only, $\rho = -0.25$, $N = 106$, $p = 0.012$; females only, $\rho = -0.20$, $N = 79$, $p = 0.074$). Moreover, those males (and again not females) who launched unprovoked attacks on their opponents had significantly higher narcissism scores than those who did not (Mann–Whitney U-test: all data, $Z = 2.23$, $N = 46,137$, $p = 0.025$; males, $Z = 2.09$, $N = 33,72$, $p = 0.037$; females, $Z = 0.92$, $N = 13,65$, $p = 0.36$; see figure 3).

4. DISCUSSION

In support of the theoretical link between overconfidence and war (Blainey 1973; Johnson 2004; Van Evera 1999), we found that: (i) subjects in a wargame were overconfident about their expectations of success in conflict; (ii) those who were more overconfident were more likely to make unprovoked attacks; (iii) overconfidence and unprovoked attacks were more pronounced among males than females; and (iv) overconfidence or unprovoked attacks were not correlated with testosterone. We also found that narcissism scores predicted both overconfidence and unprovoked attacks among males.

There are a number of possible confounding factors in our study. For example, any influence of testosterone may have been masked by other factors known to mediate the impact of this hormone on the body, such as individual variation in the androgen receptor gene, or in androgen receptor density and distribution in key neuro-anatomical

structures (Manning *et al.* 2003). Another unknown is whether males in our sample had more experience with similar kinds of tasks and/or were more engaged in such tasks, which may have increased their perceived level of confidence. In a prior study, some of our research group obtained information about the computer game habits of a similar population for a related wargame experiment (McDermott & Cowden 2001). In that study, there was no difference between the number of hours men and women had played computer games in the past, or the number of hours they currently spent playing such games. However, the type of game being played differed, such that women reported preferences for games such as Pac Man and Tetrus, while men preferred games like Mortal Kombat. Recent research suggests that in the new 'gamer generation', computer games are increasingly attracting both males and females (Beck & Wade 2004). In Beck & Wade's study of 2500 American business professionals, females accounted for 36% of frequent gamers, 67% of moderate gamers, and 77% of non-gamers. The key difference was again in the preferred types of games. Men favoured strategy, sports, racing, and action games, females favoured cerebral arcade, quiz, and puzzle games. The authors concluded that: 'What is clear is that the game world is not, as so many assume, exclusively male; that female participation continues to increase; and that gender-role behaviour is more nuanced than nongamers tend to expect' (Beck & Wade 2004: 51).

Even if there are gender differences in experience with computer games (or indeed any other cultural explanation for a gender difference in behaviour), our key finding is that males *were* overconfident; and males who were more overconfident were *more likely* to launch wars. This remains a concern irrespective of its origin: overconfidence among decision-makers may increase the chance and/or costs of war because it leads to inflated estimates of success—not necessarily of winning outright, but of likely performance, the costs involved, vulnerability to risk, and the ability to control events if things go badly (Johnson 2004).

Does this lead to useful predictions for the real world? International conflict is constant but war is not, so any plausible cause of war must exhibit *variation* to explain times of peace and times of war. Positive illusions are compelling as a cause of war because they are known to vary with specific factors. First, they vary with mental states. They are virtually absent among the depressed (a phenomenon known as 'depressive realism'), and are hugely exaggerated among those suffering from extreme narcissism or mania (a trait much more common among twentieth century leaders than in the population at large, Taylor 1989; Ludwig 2002). Second, positive illusions vary with context. They are greater, for example, in situations of ambiguity, low feedback, and where events are difficult to verify. Some researchers suggest that such contextual factors can explain 100% of the variation in positive illusions (Taylor & Armor 1996; Taylor *et al.* 2003). Third, while common to all cultures, positive illusions are relatively higher among western (especially American) populations than eastern populations (Armor & Taylor 1998; Sedikides *et al.* 2003). Fourth, the influence of positive illusions on policy outcomes varies with regime type and the decision-making process. In sum, a number of specific situations may conspire to

exacerbate or nullify positive illusions. A fairly explicit theory can therefore be constructed to derive predictions for when we can expect to see positive illusions in real world decision-making, when they are likely to contribute to causing war and, potentially, how to reduce them (Johnson 2004).

Of course, the direct applicability of our findings to the real world is heavily limited by a number of features, including the artificial laboratory setting, a situation and environment that differs markedly from real world political decision-making, anonymity, dyadic two-player interactions, a small number of iterations, and decisions that were hardly a matter of life and death. Subjects may play games to win, and can take greater risks than they would in real life (Beck & Wade 2004). In addition, the data in this study come from subjects who tended to be of a certain age, educational attainment and cultural background. The effects described here may therefore be different among people of different social, demographic and cultural backgrounds. However, these subjects were not drawn from an unduly narrow demographic base. The point of this experiment was to take a step towards testing for positive illusions in a situation *more* like war than has been attempted until now. This study reports on actual behaviour, and not merely a self-report of attitudes or hypothetical responses. Further, recent work indicates that simulated and real behaviour follow similar pathways in the brain (Jeannerod & Decety 1995). Many neurological and physiological pathways influencing decision-making and behaviour are therefore likely to be the same in the laboratory and in the halls of government, even if the magnitude of effect is very different. There is little reason to suppose that the *direction* of effects is wrong. Indeed, the pressures of limited time, high stakes, and stress in typical crisis decision-making among political or military leaders may exacerbate the effects of psychological biases rather than eliminate them (Nicholson 1992; McDermott 2004; Rosen 2004). It is also worth noting that wargames are not just games. Militaries across the world expend a large amount of time and resources conducting and running wargames to train and prepare their forces for real events. Some are very simple. In the 1980s the US Army used a modified commercial Atari game 'Battlezone' for gunnery training, and the US Marines have more recently used a version of 'Doom' to train for urban combat (Handley 2003). Others are vastly more far-reaching. The Millennium Challenge war game run by the US Department of Defense in 2002, for example, was a key stage in examining scenarios for the invasion of Iraq in 2003 and cost $250 million (US Joint Forces Command 2002). Since militaries are often concerned with how wargames represent real war, there is a significant need to understand human biology and behaviour in wargames, whether or not they also reflect real war.

Three examples illustrate how specific aspects of our findings may or may not map onto the real world and raise some interesting hypotheses that could be tested in the future. First, although we found that self-rankings were not related to actual rankings, in the real world one might expect these variables to be inextricably linked. Small countries like Liechtenstein, for example, do not usually attack large countries like Russia. There are, clearly, limits on positive illusions beyond which an inconsistency between real and perceived power would be

untenable. Nevertheless, there is good evidence that a mismatch between a state's *real* power and a state's *perceived* power is quite common, and that this often leads to war (Blainey 1973; Van Evera 1999; Johnson 2004). Furthermore, materially weaker sides often do fight and defeat more powerful opponents, such as the Mujahideen victory over the Soviet Union in Afghanistan or the Viet Cong's victory over the United States in Vietnam (Paul 1994; Arreguín-Toft 2005), along with countless examples of smaller-scale battles in which weaker units attacked and defeated much stronger sides (Johnson *et al.* 2002). So there is, in fact, considerable room for positive illusions of capability to go uncorrected and even apparently confirmed. One can always envision a conflict to turn out like Henry V at Agincourt rather than like Custer at the Little Big Horn. There is therefore no reason to expect a perfect correlation of real and perceived capability in either wargames or the real world.

Second, extrapolating experimental work to real world situations may be complicated by cultural variation. For example, southerners in the United States have been found to exhibit greater anger and testosterone levels in response to insults than northerners (Nisbett & Cohen 1996), and many historians have noted the striking cultural differences in the tactics, behaviour, and decision-making that characterized these two cultures in the American Civil War (e.g. McWhiney & Jamieson 1982). In our case, however, such cultural variation would only lead to conservative conclusions: If our population is 'northern', then we may expect to observe even *more* extreme behaviour in a southern population.

Third, the behaviour of our (mostly young) experimental subjects might be expected to differ from the behaviour of our (mostly old) state leaders and decision-makers. Age in our sample did not predict decisions for war (Mann–Whitney U-test: all data, $Z=0.75$, $N=46,137$, $p=0.45$; males, $Z=0.49$, $N=33,72$, $p=0.62$; females, $Z=1.64$, $N=13,65$, $p=0.10$; no subjects were excluded). However, the relationship between age and decisions for war has only recently been studied and remains unclear. One study suggests that political leaders with shorter tenures in office (and by implication *younger*) are more likely to attract military challenges than long-tenured men, and this makes democracies more likely to be challenged because of their restrictions on term limits (Gelpi & Greico 2001). Another recent study by some of our research group found regime type to be important, but in democracies *older* men tended to initiate violence more often (Horowitz *et al.* 2005; the reverse pattern was true in autocracies, where leaders have more individual power). The effect of age on decisions for war remains an empirical question.

While there is copious circumstantial and anecdotal evidence linking overconfidence and war, there has been no direct evidence that people exhibit positive illusions in decisions specifically relating to conflict, nor evidence that having positive illusions increases the probability of war. This study is a first step in that direction. Scholarship on the causes of war, which is founded on assumptions about human nature dating to Thucydides, Hobbes and Rousseau, may be usefully informed by modern empirical data on our biological and psychological predispositions towards conflict, and their proximate and evolutionary origins (Wrangham & Peterson 1996; Wilson 1999; Pinker 2002; Johnson 2004; Rosen 2004; Thayer 2004; Sagarin & Taylor in press). It is hard to ignore the gathering trend: as Nobel Laureate Daniel Kahneman noted recently, 'the bottom line is that all the biases in judgment that have been identified in the last 15 years tend to bias decision-making toward the hawkish side' (Shea 2004).

REFERENCES

Armor, D. A. & Taylor, S. E. 1998 Situated optimism: specific outcome expectancies and self-regulation. *Adv. Exp. Soc. Psychol.* **30**, 309–379.

Arreguín-Toft, A. 2005 *How the weak win wars: a theory of asymmetric conflict*. Cambridge, UK: Cambridge University Press.

Baumeister, R. F. & Boden, J. M. 1998 Aggression and the self: high self-esteem, low selfcontrol, and ego threat. In *Human aggression: theories, research, and implications for social policy* (ed. R. G. Geen & E. Donnerstein), pp. 111–137. San Diego, CA: Academic Press.

Beck, J. C. & Wade, M. 2004 *Got game: how the gamer generation is reshaping business forever*. Boston, MA: Harvard Business School Press.

Blainey, G. A. 1973 *The causes of war*. New York: Free Press.

Burnham, T., Chapman, J., Gray, P., McIntyre, M., Lipson, S. & Ellison, P. 2003 Men in committed, romantic relationships have lower testosterone. *Horm. Behav.* **44**, 119–122. (doi:10.1016/S0018-506X(03)00125-9)

Fearon, J. D. 1995 Rationalist explanations for war. *Int. Org.* **49**, 379–414.

Ganguly, S. 2001 *Conflict unending: India–Pakistan tensions since 1947*. New Delhi: Oxford University Press.

Gelpi, C. & Greico, J. 2001 Attracting trouble: democracy, leadership tenure and the targeting of militarized challenges, 1981–1992. *J. Conflict Resolut.* **45**, 794–817.

Gillham, J. E. (ed.) 2000 *The science of optimism and hope: research essays in honor of Martin E. P. Seligman*. Radnor, PA: Templeton Foundation.

Granger, D., Schwartz, E., Booth, A. & Arentz, M. 1999 Salivary testosterone determination in studies of childhood health and development. *Horm. Behav.* **35**, 18–27. (doi:10.1006/hbeh.1998.1492)

Gray, P., Campbell, B., Marlowe, F., Lipson, S. & Ellison, P. 2004 Social variables predict between-subject but not day-to-day variation in the testosterone of US men. *Psychoneuroendocrinology* **29**, 1153–1162. (doi:10.1016/j.psyneuen.2004.01.008)

Handley, R. 2003 Modeling and simulation: Hollywood fulfills military needs. *Advanced Imaging* (March issue). See http://www.advancedimagingpro.com/publication/article.jsp?pubId=1&id=118.

Haselton, M. G. & Nettle, D. 2006 The paranoid optimist: an integrative evolutionary model of cognitive biases. *Pers. Soc. Psychol. Rev.* **10**, 47–66. (doi:10.1207/s15327957pspr1001_3)

Horowitz, M., McDermott, R. & Stam, A. 2005 Leader age, regime type and violence. *J. Conflict Resolut.* **49**, 661–685. (doi:10.1177/0022002705279469)

Janis, I. L. 1972 *Victims of groupthink: psychological studies of policy decisions and fiascoes*. Boston, MA: Houghton Mifflin.

Jeannerod, M. & Decety, J. 1995 Mental motor imagery: a window into the representational stages of action. *Curr. Opin. Neurobiol.* **5**, 727–732. (doi:10.1016/0959-4388(95)80099-9)

Jervis, R. 1976 *Perception and misperception in international politics*. Princeton, NJ: Princeton University Press.

Johnson, D. D. P. 2004 *Overconfidence and war: the havoc and glory of positive illusions*. Cambridge, MA: Harvard University Press.

Johnson, D. D. P. & Tierney, D. R. 2006 *Failing to win: perceptions of victory and defeat in international politics*. Cambridge, MA: Harvard University Press.

Johnson, D. D. P., Wrangham, R. W. & Rosen, S. P. 2002 Is military incompetence adaptive? An empirical test with risk–raking behaviour in modern warfare. *Evolut. Hum. Behav.* **23**, 245–264. (doi:10.1016/S1090-5138(01)00087-3)

Lebow, R. N. 1981 *Between peace and war: the nature of international crisis*. Baltimore, MD: John Hopkins University Press.

LeShan, L. 2002 *The psychology of war: comprehending its mystique and its madness*. New York, NY: Helios.

Lipson, S. & Ellison, P. 1989 Development of protocols for the application of salivary steroid analyses to field conditions. *Am. J. Hum. Biol.* **1**, 249–255. (doi:10.1002/ajhb.1310010304)

Ludwig, A. M. 2002 *King of the mountain: the nature of political leadership*. Lexington, KY: University Press of Kentucky.

Manning, J. T., Bundred, P. E., Newton, D. J. & Flanagan, B. F. 2003 The second to fourth digit ratio and variation in the androgen receptor gene. *Evolut. Hum. Behav.* **24**, 399–405. (doi:10.1016/S1090-5138(03)00052-7)

Mazur, A. & Booth, A. 1998 Testosterone and dominance in men. *Behav. Brain Sci.* **21**, 353–397. (doi:10.1017/S0140525X98001228)

McDermott, R. 2004 *Political psychology in international relations*. Ann Arbor, MI: University of Michigan Press.

McDermott, R. & Cowden, J. 2001 The effects of uncertainty and sex in a simulated crisis game. *Int. Interact.* **27**, 353–380.

McDermott, R., Cowden, J. Submitted. Hostile communication in a crisis simulation game.

McWhiney, G. & Jamieson, P. D. 1982 *Attack and die: civil war military tactics and the southern heritage*. Alabama, AL: University of Alabama Press.

Muller, M. N. & Wrangham, R. W. 2001 The reproductive ecology of male hominoids. In *Reproductive ecology and human evolution* (ed. P. Ellison), pp. 397–427. New York, NY: Aldine.

Nettle, D. 2004 Adaptive illusions: optimism, control and human rationality. In *Emotion, evolution and rationality* (ed. D. Evans & P. Cruse). Oxford, UK: Oxford University Press.

Nicholson, M. 1992 *Rationality and the analysis of international conflict*. Cambridge, UK: Cambridge University Press.

Nisbett, R. E. & Cohen, D. 1996 *Culture of honor: the psychology of violence in the south*. Denver, CO: Westview Press.

Paul, T. V. 1994 *Asymmetric conflicts: war initiation by weaker powers*. Cambridge, UK: Cambridge University Press.

Peterson, C. 2000 The future of optimism. *Am. Psychol.* **55**, 44–55. (doi:10.1037/0003-066X.55.1.44)

Pinker, S. 2002 *The blank slate: the modern denial of human nature*. New York, NY: Penguin Putnam.

Raskin, R. & Terry, H. 1988 A principal-components analysis of the Narcissistic Personality Inventory and further evidence of its construct validation. *J. Pers. Soc. Psychol.* **54**, 890–902. (doi:10.1037/0022-3514.54.5.890)

Rosen, S. P. 2004 *War and human nature*. Princeton, NJ: Princeton University Press.

Sagarin, R. & Taylor, T. (eds) In press. *Darwinian security: perspectives from ecology and evolution*. Berkeley, CA: University of California Press.

Sedikides, C., Gaertner, L. & Toguchi, Y. 2003 Pancultural self-enhancement. *J. Pers. Soc. Psychol.* **84**, 60–79.

Shea, C. 2004 The power of positive illusions. *The Boston Globe*, 26 September 2004. See http://www.boston.com/news/globe/ideas/articles/2004/09/26/the_power_of_positive_illusions.

Stoessinger, J. G. 1998 *Why nations go to war*. New York, NY: St Martin's.

Taylor, S. E. 1989 *Positive illusions: creative self-deception and the healthy mind*. New York, NY: Basic Books.

Taylor, S. E. & Armor, D. A. 1996 Positive illusions and coping with adversity. *J. Pers.* **64**, 873–898. (doi:10.1111/j.1467-6494.1996.tb00947.x)

Taylor, S. E. & Brown, J. D. 1988 Illusion and well-being: a social psychological perspective on mental health. *Psychol. Bull.* **103**, 193–210. (doi:10.1037/0033-2909.103.2.193)

Taylor, S. E. & Brown, J. D. 1994 Positive illusions and well-being revisited: separating fact from fiction. *Psychol. Bull.* **116**, 21–27. (doi:10.1037/0033-2909.116.1.21)

Taylor, S. E., Lerner, J. S., Sherman, D. K., Sage, R. M. & McDowell, N. K. 2003 Portrait of the self-enhancer: well adjusted and well liked or maladjusted and friendless. *Pers. Process Indiv. Differ.* **84**, 165–176.

Thayer, B. A. 2004 *Darwin and international relations: on the evolutionary origins of war and ethnic conflict*. Lexington, KY: University Press of Kentucky.

Tiger, L. 1979 *Optimism: the biology of hope*. New York, NY: Simon & Schuster.

Trivers, R. L. 2000 The elements of a scientific theory of self-deception. *Ann. NY Acad. Sci.* **907**, 114–131.

US Joint Forces Command 2002. See http://www.jfcom.mil/about/experiments/mc02.htm.

Van Evera, S. 1999 *Causes of war*. Ithaca, NY: Cornell University Press.

Van Evera, S. 2003 Why states believe foolish ideas: nonself-evaluation by states and societies. In *Perspectives on structural realism* (ed. A. K. Hanami), pp. 163–198. New York, NY: Palgrave Macmillan.

Wilson, E. O. 1999 *Consilience: the unity of knowledge*. London, UK: Abacus.

Wrangham, R. 1999 Is military incompetence adaptive? *Evolut. Hum. Behav.* **20**, 3–17. (doi:10.1016/S1090-5138(98)00040-3)

Wrangham, R. & Peterson, D. 1996 *Demonic males: apes and the origins of human violence*. London, UK: Bloomsbury.