Available online at www.sciencedirect.com

ScienceDirect

Hormones and Behavior 51 (2007) 373–378

Hormones and Behavior

www.elsevier.com/locate/yhbeh

Testosterone levels and mental rotation performance in Chinese men

Chi-Fu Jeffrey Yang [a], Carole K. Hooven [b], Matthew Boynes [c], Peter B. Gray [d], Harrison G. Pope Jr. [c,*]

[a] Harvard Medical School, Boston, MA 02115, USA
[b] Department of Anthropology, Harvard University, Cambridge, MA 02138, USA
[c] Biological Psychiatry Laboratory, McLean Hospital/Harvard Medical School, Belmont, MA 02478, USA
[d] Department of Anthropology, University of Nevada Las Vegas, Las Vegas, NV 89154, USA

Received 5 September 2006; revised 1 December 2006; accepted 4 December 2006
Available online 22 December 2006

Abstract

Males achieve markedly higher scores than females on mental rotation tests (MRTs). Therefore, it might be hypothesized that, within groups of males, testosterone levels modulate MRT performance. However, studies of this relationship have yielded inconsistent results. Notably, a recent study of 28 American men, using the computerized Shepard and Metzler MRT (SM), found significant associations between salivary testosterone levels and the intercepts of the functions relating response time and error rate to the angular disparity between comparison objects. Conversely, a study of 35 British men, using the same methodology, found no such associations. We attempted a cross-cultural replication of these studies, in which we obtained salivary testosterone levels, together with performance measures on the SM, from 92 heterosexual right-handed men, aged 21–38, in Beijing, China. We hypothesized that Chinese men might perform more slowly and carefully than Western men on this test (which imposes no time limitations), but that associations of testosterone levels with performance, if real, should nevertheless be detectable across cultures. We found that the Chinese men indeed displayed significantly longer response times than the American men, although the Chinese men were equally accurate. Interestingly, testosterone was significantly associated with the *slope* of the response time function in Chinese men, whereas the earlier American study had found that testosterone was associated with the *intercept*, but not the slope, of this function. These observations suggest that differing cultural values regarding speed and accuracy may influence MRT performance — and that these values must be considered in future studies of testosterone and MRT measures.

Keywords: Spatial ability; Mental rotation test; China; Testosterone; Male; Cross-cultural

Introduction

In studies comparing the sexes on cognitive abilities, the most consistent differences have been observed on mental rotation tests (MRTs), where males typically score 0.7 to 0.94 standard deviations above females — a finding reported across cultures (Voyer et al., 1995; Halpern and Tan, 2001; Oosthuizen, 1991). This difference in performance has been suggested to be at least partly attributable to *organizational* effects of fetal testosterone – namely, permanent and early established effects of testosterone on brain structure (Liben et al., 2002; Hines, 2004). However, it also seems possible that *activational* effects of current testosterone levels may modulate MRT performance, over and above any possible contribution from the organizational effects of fetal testosterone levels. Since MRTs appear to be the most powerful instruments for detecting differences between the sexes in cognition, then by inference they might be the best tests to detect possible activational effects of testosterone within males. But here findings have been inconsistent, possibly due in part to differences among the subject populations and the specific MRTs used (see Liben et al., 2002; Halari et al., 2005 for reviews). Of the various MRTs used in these studies, the most common is the Vandenberg and Kuse (VK) (Vandenberg and Kuse, 1978), a paper-and-pencil test yielding a single accuracy score. The VK is an adaptation of the original MRT of Shepard and Metzler (SM) (Shepard and

* Corresponding author. Fax: +1 617 855 3585.
E-mail address: hpope@mclean.harvard.edu (H.G. Pope).

0018-506X/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.yhbeh.2006.12.005

Metzler, 1971). The SM, described in more detail below, is a computerized test that measures response time and error rate as a function of the degree of angular disparity between the two stimulus objects. Throughout the literature, the VK and SM are the most widely used MRTs (Voyer et al., 1995).

We reviewed all studies assessing the relationship between mental rotation performance and testosterone levels in groups of normal men, and found a total of nine — seven using the VK, and two (Hooven et al., 2004; Falter et al., 2006) using the SM. Three found significant positive associations: Silverman et al. (1999), using salivary testosterone; Gordon and Lee (1986), using serum total testosterone; and Hooven et al. (2004), using salivary testosterone. One study found a significant negative association between salivary testosterone and MRT scores (Moffat and Hampson, 1996). The remaining five studies found no significant association between MRT scores and measures of salivary testosterone (Gouchie and Kimura, 1991; Falter et al., 2006), serum total testosterone (Alexander et al., 1998), plasma total testosterone (Kampen and Sherwin, 1996), or "venous" total and free testosterone (Halari et al., 2005).

The inconsistencies in these studies may reflect methodological limitations, such as small sample sizes (a median sample size of 33 normal men, with only 59 men in even the largest study (Silverman et al., 1999)). Another possible limitation is reliance on serum or plasma total testosterone levels, which may not reflect biologically available testosterone as accurately as salivary testosterone levels (Vining and McGinley, 1987; Lac, 2001). Specifically, sex hormone binding globulin (SHBG) may not be transported readily from plasma into saliva, and hence salivary testosterone levels may better reflect the unbound, biologically available fraction of the hormone. Other limitations on existing studies include differences in the timing of testosterone levels relative to MRT testing time; limitations in the accuracy of testosterone assay methods (leading perhaps to type II errors); and even the presence of female investigators in the room at times of testosterone collection — an influence shown capable of quickly altering testosterone levels (Roney et al., 2003).

Another possible methodological limitation, proposed by Hooven et al. (2004), is that simple raw MRT scores, such as those typically generated by the VK, are too crude to capture specific aspects of performance, and thus fail to detect the true relationship of testosterone levels to mental rotation ability. Performance on an MRT, like all cognitive tasks, requires several different cognitive processes, such as visual encoding, rotation, and decision-making (Karadi et al., 2001) — and circulating testosterone might be related to some of these individual processes, but not to others (Hooven et al., 2004). For example, males with higher testosterone levels might be quicker or more accurate in their ability to visually encode the objects, to imagine them as three-dimensional objects in space, or to make decisions under pressure. Clearly, it would be desirable to design studies to isolate each of the above cognitive steps, especially the rotation aspect of the task. The computerized SM, unlike the VK, permits an analysis of rotation, since it measures both error rate (ER) and response time (RT) for individual trials; thus, one can calculate the slope and intercept of the functions relating ER and RT performance to the degree of angular disparity between the comparison objects. It is thought that the slopes of these functions index the rotation process itself, whereas the intercepts index the contribution of all other processes used to perform the task (Hooven et al., 2004).

Two recent studies have employed this methodology. Hooven et al. (2004) administered the SM to 28 normal American men, and found that salivary testosterone levels were negatively correlated with ER and RT (in other words, higher testosterone predicted more accurate and faster responses) — but testosterone levels were associated only with the intercepts of the rotation functions and not the slopes. This surprising finding suggested to the authors that testosterone might affect performance on MRTs through its relationship to a non-rotation process, such as decision-making. By contrast, Falter et al. (2006), in the course of a study using several cognitive tests, obtained SM results in 35 normal British men. Using a somewhat easier version of the SM, but employing a method of analysis identical to Hooven et al., these investigators found no association of testosterone with either the slope or intercept function.

In light of these conflicting findings, we performed a study in which we obtained salivary testosterone levels and SM performance measures from 92 men in Beijing, China, using the SM version of Hooven et al. (2004). We then analyzed these data using both a conventional analysis (looking at the simple association of testosterone levels with overall performance and overall accuracy) and a second, more detailed analysis using the methods of the Hooven et al. and Falter et al. studies just described.

This study is the first, to our knowledge, to examine the association between testosterone and MRT performance in a non-Western population. We felt that it would be particularly interesting to study Chinese men in this respect, because their different cultural values might affect their approach to MRT testing. Specifically, the Confucian philosophical tradition emphasizes accuracy and thoroughness over pure speed. For example, in the *Analects* of Confucius there is a saying still commonly used today in China – *yusu buda* (欲速不達) – which cautions that rushing to achieve a goal will lead to failure (Li, 1997). By contrast, it has been argued that American schooling, with its frequent use of timed tests, reinforces the value of speed (Nell, 2000). These impressions are supported by one recent study of simple, choice, and odd-man-out reaction time (Jensen and Whang, 1993) that found slower reaction times in Chinese–Americans as opposed to Anglo-Americans. The authors of this study conjectured that the Chinese–American participants might be sacrificing speed for accuracy of response.

For these reasons, we speculated that Chinese participants might perform more slowly, but with fewer errors, than their American counterparts. However, we hypothesized that if testosterone has a true activational influence on cognition, one should find an association between testosterone levels and MRT performance across cultures in spite of these differences. In other words, the intercepts and slopes of the response functions might differ between cultures, but activational effects of testosterone on MRT performance, if real, would still be evident in the form of significant associations between current testosterone levels and specific performance measures.