Neuroscience and Biobehavioral Reviews 184 (2026) 106575

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www elsevier com/locate/neubiorev

No relationship between testosterone and risk aversion: A meta-analytic review

Irene Sánchez Rodríguez [a], Luca Bailo [a], Folco Panizza [a], Emiliano Ricciardi [a], Francesco Bossi [b,*]

[a] Molecular Mind Lab IMT School for Advanced Studies Lucca Piazza San Francesco 19 Lucca 55100 Italy
[b] Department of Information Engineering University of Pisa Largo Lucio Lazzarino 1 Pisa 56122 Italy

ARTICLE INFO

Keywords:
Risk aversion
Risk preferences
Testosterone

ABSTRACT

The association between testosterone and risk-taking behavior has been widely investigated across behavioral economics, neuroendocrinology, and social neuroscience, but empirical results remain inconsistent. To clarify this relationship, we conducted a multilevel random-effects meta-analysis of 52 empirical studies (94 independent effect sizes; total N = 17,340), the most comprehensive so far, examining correlations between testosterone levels or manipulations and risk preferences across diverse paradigms. The aggregated effect was statistically null (r = –0.0021, 95 CI [–0.0431, 0.0389], p = .919), indicating no reliable link between testosterone and risk-taking. Publication bias diagnostics (trim-and-fill and fail-safe N) suggested that this null effect is not driven by selective reporting. Meta-regressions revealed significant heterogeneity across testosterone measurement type. Moreover, only lottery-based economic tasks showed a modest positive association, whereas other paradigms (e.g., BART, IGT, self-report) did not. A separate meta-analysis of sex differences found no moderating effect, suggesting that testosterone-risk correlations are not reliably stronger in males than females. Overall, the evidence challenges the notion that testosterone provides a general hormonal basis for human risk preferences. Instead, findings support a biopsychosocial framework in which "risk taking" reflects the interaction of task demands, cognitive–affective processes, and situational context, with endocrine effects appearing narrow, context-dependent, and method-specific. Future work should employ preregistered, multi-measure designs and direct endocrine assays to test mechanistic pathways more precisely.

Contents

1. Introduction 2
2. Method 3
 2.1. Literature search methodology 3
 2.2. Selection criteria 3
 2.3. Meta-analytic methods 3
  2.3.1. Meta-regression 4
  2.3.2. Sex differences 5
3. Results 5
 3.1. Meta-analysis - overall effect 5
 3.2. Trim and fill method 6
 3.3. Meta-regression - moderators 6
 3.4. Meta-regression - sex differences 6
4. Discussion 6
Acknowledgment 9
Appendix A. Supplementary Table A.1 9
Appendix B. Supplementary data 9

* Corresponding author.
E-mail address: francesco.bossi@ing.unipi.it (F. Bossi).

https://doi.org/10.1016/j.neubiorev.2026.106575
Received 13 October 2025; Received in revised form 12 December 2025; Accepted 23 January 2026
Available online 2 February 2026

References 10
Additional articles included in the meta-analysis 11

1. Introduction

Risk aversion, the tendency to prefer options with lower outcome variability over riskier but potentially more rewarding alternatives, shapes choices that range from everyday health preferences (e.g., smoking, driving, vaccination uptake) to high-stakes financial and political decisions (Lakdawalla et al., 2020; Guiso and Paiella, 2008). Individuals differ in their preferences under uncertainty: some prioritize higher potential rewards despite elevated risks, whereas others opt for safer but less rewarding outcomes. Understanding why people differ in risk preferences is, therefore, central to behavioral science, economics, and public health policy (Sapienza et al., 2009a; Stanton et al., 2011a).

Multiple factors contribute to individual differences in risk aversion, including personality traits (Soane and Chmiel, 2005; Mueller et al., 2006), decision framing (Kahneman and Tversky, 1979; Von Neumann and Morgenstern, 1944), as well as social-emotional context (Loewenstein et al., 2001). Among these, sex differences stand out as one of the most robust findings: on average, males tend to be less risk-averse than females (Byrnes et al., 1999).

To explain this gap, hypotheses have been proposed ranging from social and cultural perspectives to biological ones. Social-cultural explanations emphasize gender differences in roles and expectations, educational and socialization pathways, and differences in context sensitivity when assessing risks, underscoring the malleability of risk preferences and the influence of gendered norms (Eckel and Grossman, 2008; Sarin and Wieland, 2016; Rai and Kimmel, 2015; Friedl et al., 2020). By contrast, biological hypotheses, especially those focusing on sex-linked hormones, have received greater empirical attention, with numerous studies testing whether between-sex differences in androgens (e.g., testosterone) are associated with systematic differences in risk preferences (Apicella et al., 2008; Zethraeus et al., 2009a).

Testosterone, which differs on average between males and females, has been proposed as a key driver of sex differences in risk. However, the empirical record is mixed. So far, no solid explanation has emerged to explain hormonal differences in risk aversion. Some studies reported a significant correlation between testosterone and risk aversion, observing that higher baseline testosterone levels have been associated with increased financial risk taking (Apicella et al., 2008), while others reported that elevated testosterone levels attenuate or eliminate gender differences in risk aversion (Sapienza et al., 2009a). In contrast, several studies reported no relationship between risk preferences and testosterone levels (Boksem et al., 2013b; Derntl et al., 2014b; Zethraeus et al., 2009a). This pattern raises the question of whether testosterone provides a general explanatory route to risk preferences or whether observed links are context-bound.

In response to inconsistent single-hormone findings, more complex models have emerged that consider testosterone alongside other factors. Chief among these is the dual hormone-hypothesis, which proposes that testosterone's behavioral effects depend on concurrent cortisol levels (Carre and McCormick, 2008). This perspective aligns with evidence for stress-related shifts in financial risk taking (Kandasamy et al., 2014) and small sample reports of testosterone and cortisol associations on laboratory tasks (Mehta and Prasad, 2015). Yet, the empirical base remains thin and methodologically heterogeneous, limiting the feasibility of cumulative inference.

The discrepancies in these results, both for the single- and dual-hormone hypotheses, pose serious obstacles to synthesizing the evidence and clarifying testosterone's role in sex differences in risk aversion. A critical source of heterogeneity across the literature is how both risk and hormone measures are operationalized. Risk preferences have been elicited with paradigms that recruit different computations and affective processes. For instance, the Balloon Analogue Risk Task (BART) captures impulsive decision making under uncertainty (Strobel et al., 2001), the Iowa Gambling Task (IGT) emphasizes feedback-based learning and loss updating (Tversky and Kahneman, 1994), and the Holt & Laury lottery task targets the evaluation of long-run expected utility under known probabilities (Holt and Laury, 2002). Comparable heterogeneity appears on the endocrine side, with studies relying on divergent approaches. These include the active administration of testosterone; direct biochemical quantification of hormone concentrations (e.g., saliva or blood assays) (Ramirez et al., 2003); and indirect proxies based on morphological traits intended to index prenatal androgen exposure, such as the 2D:4D digit ratio (Mueller et al., 2006) and face masculinity (Jackson et al., 2005). Importantly, large and well-powered work has questioned the validity of some morphology based proxies, most notably 2D:4D, as predictors of economic preferences, including risk (Alonso et al., 2018; Neyse et al., 2021). These methodological differences, on both the behavioral and endocrine sides, help explain why findings can diverge across studies that nominally address the same question.

Building on this background, we ask a precise question: is testosterone robustly associated with human risk preferences? We aim to perform the most comprehensive quantitative synthesis of studies investigating the impact of testosterone to date and to understand whether it can explain risk preferences by itself, or whether a broader biopsychosocial view better fits the evidence. We test this by bringing together studies that use different tasks and different ways of measuring hormones, and by asking whether the pattern of results tracks what the tasks actually measure and how hormones are assessed. Because sex differences motivate much of this work, we examine sex as a moderator of the testosterone–risk link. That is, whether the hormone-behavior association differs for males and females rather than estimating mean sex differences per se.

While previous reviews have already examined the hormone-risk relationship (e.g., Apicella et al., 2015; Stanton, 2017) and discussed potential explanations for mixed findings, they have been narrative in nature and were limited to subsets of the literature (for example, focusing on economic or consumer decision-making and predominantly on single-hormone accounts). In contrast, our work aims to provide a comprehensive quantitative synthesis of the association between hormones and risk preferences across both endogenous and administration studies, integrating all major studies on both single- and dual- hormone hypotheses identified in our search. In addition, we evaluate evidential value using standard publication-bias diagnostics.

Although interactionist models (e.g., testosterone with cortisol) are conceptually relevant to our focus, the current evidence base is too slim and heterogeneous for a cumulative estimate; we therefore discuss these claims narratively and outline design priorities for future tests. Similarly, due to our pre-specified search date, some highly relevant works were not included in the quantitative synthesis. We discuss these more recent studies in the Discussion section as convergent evidence bearing on the questions addressed here.

Finally, our synthesis combines studies that relate endogenous baseline testosterone to risk preferences with experimental testosterone administration studies, allowing us to integrate both correlational and causal evidence. We analyze all effect sizes using multilevel random-effects models with a random intercept for each sample, which appropriately accounts for multiple risk tasks and different types of hormone indices (direct endocrine measures, morphological proxies, and pharmacological manipulations) reported within the same study.

2. Method

2.1. Literature search methodology

The search for relevant studies was carried out using an electronic search of different databases: Google Scholar, PubMed and Scopus. The following Boolean string was applied to titles, abstracts, and keywords in each database: testosterone AND (“risk attitude” OR “risk seeking” OR “risk aversion”). We deliberately chose this more specific risk terminology, rather than broader terms such as “risk taking” or “risk”, to target the economic and psychological risk-preference literature and to avoid records on health or clinical risk outcomes unrelated to choice behavior. The search was completed on 2 March 2023 with no restriction on country. Where available, database filters were set to “Humans” and meta-analyses and reviews were discarded. All records were exported, and duplicates were removed automatically and then verified manually. Full database-specific search strings, including field tags and filters for PubMed, Scopus, and Google Scholar, are provided in Supplementary Table A.1 to facilitate independent replication of the search.

Even though it did not represent the focus of this study, we also searched the literature for studies that focus on both single- and dual-hormone hypotheses. However, we could not find enough data (i.e., 7 studies) to proceed with a meta-analysis of the dual hormone hypothesis, while sufficient material was present to examine studies that investigated the role of testosterone itself.

2.2. Selection criteria

Studies were included in the current review and meta-analysis if they satisfied the following criteria:

> Studies had to report a statistical association between testosterone and risk preference outcome.
> Studies that measure or administer testosterone
> Use at least one behavioral or self-report measure of risk preference (e.g. Balloon Analogue Risk Task, Iowa Gambling Task, lottery or investment tasks, risk-taking scales)
> Studies reported sufficient data to calculate an effect size (*r d t F beta* or exact *p*); missing data obtainable from the authors.
> Peer-reviewed articles, dissertations, preregistered reports, or in-press manuscripts written in English, Spanish or Italian. Conference abstracts were included only when a full paper and statistics were available.

All studies were cross-referenced to avoid duplicates in the meta-analysis. Reviews and chapters in books were excluded, and only empirical studies were considered. We further excluded purely behavioral studies that did not measure or manipulate testosterone, studies that relied exclusively on facial masculinity or other facial metrics as proxies for testosterone.

We restricted inclusion to reports written in English, Spanish, or Italian; records in other languages were excluded at title and abstract screening. Studies with underage participants were excluded, with an exception for Stenstrom et al. (2011), whose sample had an age range from 17 to 44 years old, with a mean age of 20.9 years. We focused on non-clinical populations and excluded studies that recruited participants on the basis of diagnosed psychiatric, neurological, or endocrine disorders, or chronic pharmacological treatments directly targeting sex hormones. We did not impose minimum reporting thresholds for assay sensitivity, intra- or inter-assay coefficients of variation, or precise sampling time of day; studies were eligible as long as they included a clearly specified testosterone measure or manipulation.

Study selection proceeded in two stages. First, one author (ISR) screened all titles and abstracts against the inclusion criteria and excluded records that were clearly ineligible (e.g., no testosterone measure or manipulation, no risk outcome, non-human samples, reviews or theoretical papers). For records whose inclusion was not clear, they were evaluated by two additional authors (LB and FB). For the independent evaluation by two reviewers, an inter-rater reliability measure was computed: $\kappa =$ 91, 95 CI [0.86 0.96]. This result shows a very high inter-rater agreement. In cases of disagreement, a third author (FP) was consulted until a consensus was reached.

The selection process is represented by the following flow diagram (see Fig. 1).

2.3. Meta-analytic methods

The whole meta-analysis was carried out using RStudio software (RStudio Team, 2023). The package “compute.es” (Del Re, 2013) was used to estimate and transform effect size indices. The package “metafor” (Viechtbauer, 2010) was used to perform the meta-analysis and create the plots presented in the Results section. The effect size index we used for all outcome measures was the Pearson correlation coefficient *r*, measuring linear correlation between two sets of data. The effect size was determined based on the reported statistics, including *r*, *t*, *F*, and *p* values.

When the effect was reported as “significant, $p < .05$”, with no further information, we computed the effect size on $p = .05$ to obtain a conservative measure (this was necessary only for two effect size measures). Forty-three out of 52 studies presented a within-subjects or mixed between/within-subjects design. To account for the resulting dependency in the data, we adopted a multilevel random-effects meta-analytic model. In random-effects models, true effects are assumed to be drawn from a distribution of possible effects, accounting for both between-study and within-study variability. More specifically, we modeled a random intercept for each sample to appropriately account for the clustering of effect sizes within studies. This approach ensures that, when multiple effect sizes were reported from the same sample, for example, through different measures of risk (e.g., a self-report scale and the Balloon Analog Risk Task) or testosterone (e.g., 2D:4D ratio and salivary levels), multiple time points, or multiple doses within a study, the shared variance due to the common sample was modeled explicitly, thereby avoiding over representation of that study in the overall meta-analytic estimate.

Different measures of risk-testosterone correlation were extracted from each study. Testosterone measures/manipulation were grouped into four levels: direct administration (7 measures), blood sample (6), fingers measures (i.e., 2D:4D ratio or rel2 ratio) (51), and saliva sample (30). Risk measures were grouped into seven levels: Balloon Analog Risk Task (BART) (15), Iowa Gambling Task (IGT) (8), other behavioral (ad hoc) tasks (5), investment tasks (7), lottery tasks (Holt & Laury or derived) (37), self-report scales (20), and trading simulations (2). One effect size was computed for each of them, 94 measures in total extracted from 52 studies. For each study, we also coded basic endocrine characteristics, including the type of testosterone measure (direct endocrine measure, morphological proxy, or pharmacological administration) and the sample matrix (saliva, blood, or other). However, additional assay details such as sampling time of day, assay sensitivity, and intra- and inter-assay coefficients of variation, were too inconsistently available across studies to be used as formal inclusion criteria or as moderators in the meta-analytic models.

Within the subset of studies using direct endocrine measures (saliva or blood), almost all effect sizes were based on baseline (tonic) testosterone levels. Only a single effect size, from Stenstrom et al. (2018), used within-person change in salivary testosterone (ΔT) as the predictor of risk preferences, and this association was non-significant.

Heterogeneity across sets of outcomes was assessed using the $Q_{homogeneity}$ statistic. A statistically significant result in the test on $Q_{homogeneity}$ represents high heterogeneity across the results of different studies. It is important to remember that the $Q_{homogeneity}$ tends to be

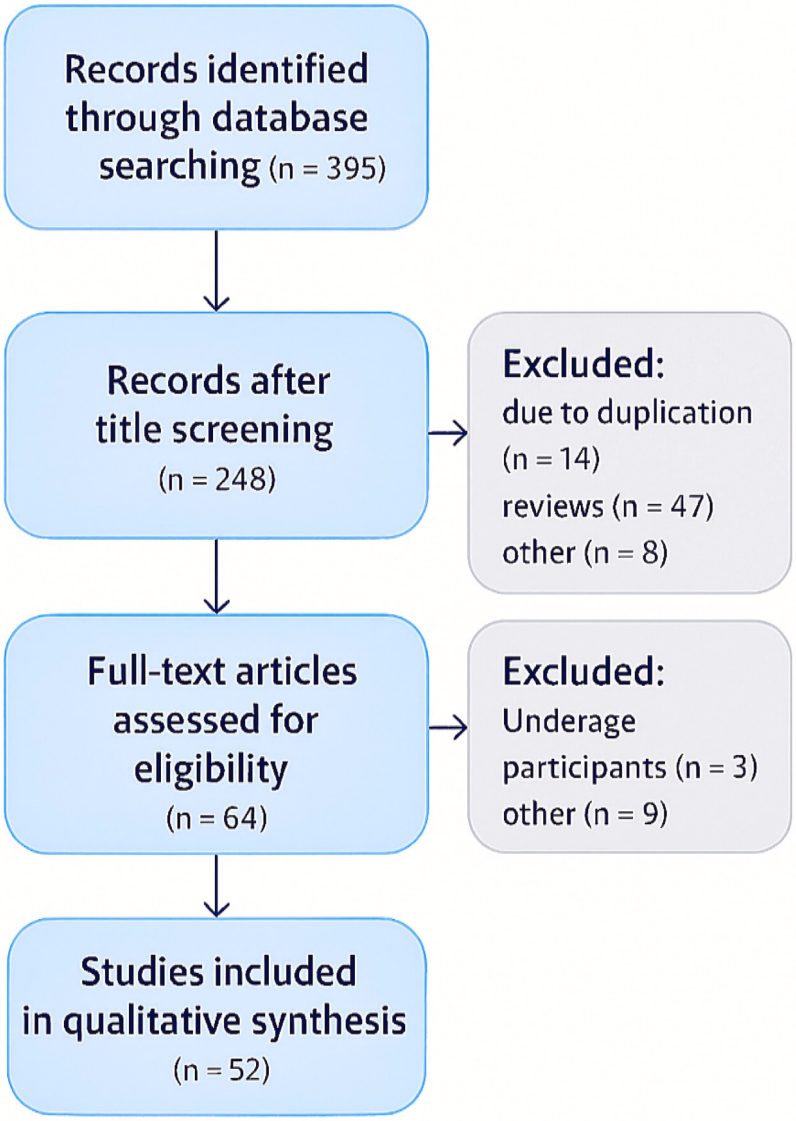

Fig. 1. PRISMA flow diagram of study selection for the meta-analysis of testosterone and risk preferences.

always significant with a large number of studies considered (as every test assessing data distribution) (Cohen et al., 2013).

The "file drawer problem" (Mullen, 1989) defines the fact that only studies reporting statistically significant results are typically published in international peer-reviewed journals, while experiments showing null results tend to be "left behind" by researchers. In order to calculate the effect of potential data censoring or publication bias on the results of the meta-analysis, the "trim and fill" method (Duval and Tweedie, 2000a,b) was used. Following this non-parametric method, a funnel plot of the effect size of each measure was designed against the standard error (on the Y axis). If no publication bias is present, this plot is expected to have the shape of a funnel, because studies with smaller sample sizes (and thus, larger standard errors) have increasingly larger variation in estimates of their effect size because random variation becomes increasingly influential, while studies with larger sample sizes have smaller variation in effect sizes. With the trim and fill procedure, studies with the highest effect size compared to their standard error (thus, out of the funnel shape), which are considered to be symmetrically unmatched by missing studies, are trimmed and their missing counterparts are "filled" as mirror images of the trimmed outcomes.

Another method typically used in meta-analyses is the Fail-safe method (Rosenthal, 1991). This method is aimed at estimating the number of studies with null results that should be added to the literature to reduce the considered effect size to "non-significantly different from zero". Typically, the fail-safe number (Nfs) is representative of the robustness of the effect size.

2.3.1. *Meta-regression*

As stated in the Introduction, the correlation between risk and testosterone was reported to be modulated by different variables (that is, the testosterone measure and the risk measure).

For this reason, a meta-regression was carried out. A meta-regression is a method aimed at investigating the influence of one or more predictors used in the included studies. The influence of both predictors (i.e., testosterone measure and risk measure) on combined effect sizes was tested by using likelihood ratio tests between nested models in a mixed fixed-random effects model. A statistically significant likelihood ratio test indicates that the difference in effect size between subsets of studies is significant. Moreover, we tested the contrasts between each level of the moderator vs. the mean value via z-tests.

The predictors tested in the meta-regression were described above and are (i) testosterone measure category (4 levels), and (ii) risk measure category (7 levels).

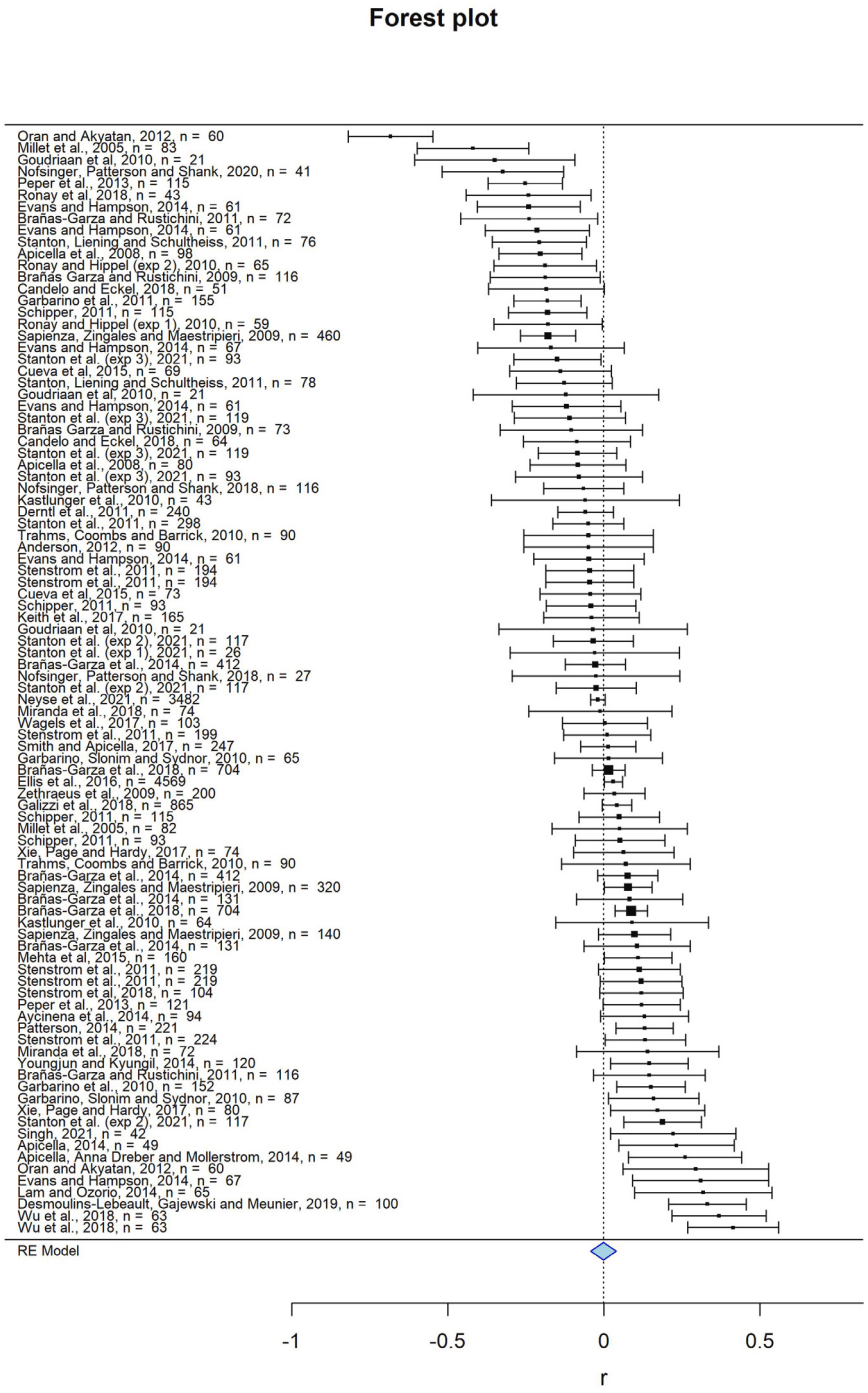

Fig. 2. The forest plot represents the effect size of each study, together with its 95 confidence interval (error bars) and sample size (represented by the squares' dimension). The mean effect size (Pearson's r = −0.0021) computed by the multilevel random-effects model is represented by the diamond in the lower part of the plot, while the diamond's width represents the 95 confidence interval.

2.3.2. Sex differences

Since previous literature showed the crucial importance of sex differences in the risk-testosterone correlation, we also tested the moderating effect of participants' sex in a separate multilevel meta-analytic model. This meta-analysis was performed separately because only a subsample of studies reported separate effect sizes for male vs. female participants, i.e., 76 observations from 37 studies. The effect of this moderator was tested with a likelihood ratio test as explained above.

In line with current guidelines on sex and gender in research (Heidari et al., 2016), we use the word "sex" to refer to male/female categories as recorded in the primary studies, which typically reflect biological or legal classification, and reserve "gender" for sociocultural roles or identities when these are explicitly measured or discussed.

3. Results

3.1. Meta-analysis - overall effect

The combined effect size of the 52 studies (94 measurements) was r = −0.0021, total N = 17340, 95 CI [−0.0431, 0.0389], z = −0.1018,

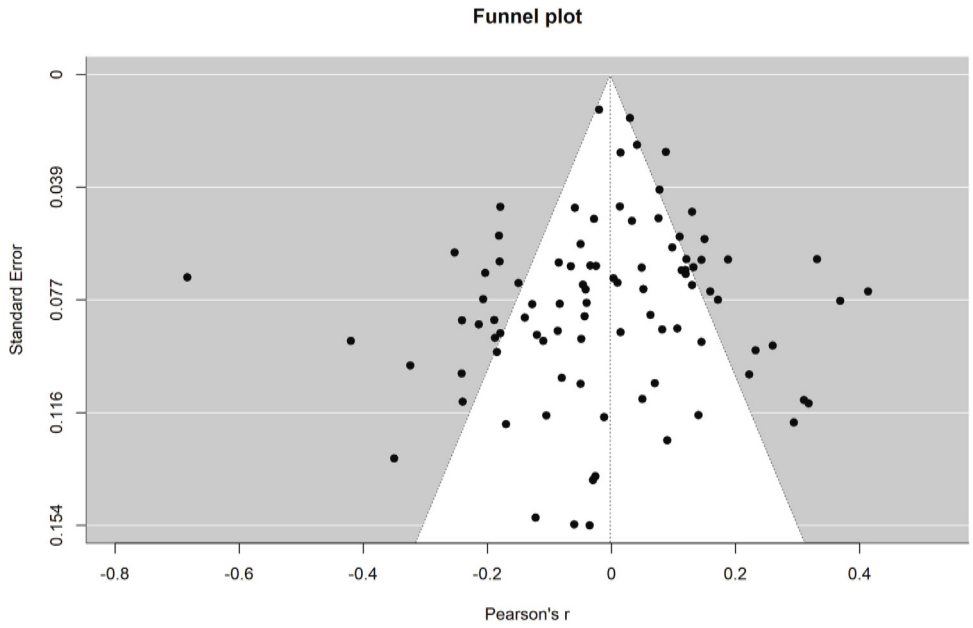

Fig. 3. The funnel plot represents included studies as black dots. In this plot the effect size is shown on the X axis and the standard error (inversely proportional to the sample size) on the Y axis. The white funnel shape represents the area where all studies should lie symmetrically if there were no publication bias, showing that studies with smaller sample sizes (and thus, larger standard errors) have increasingly larger variation in estimates of their effect size. The trim and fill method estimated 0 studies with null results subject to publication bias.

p = .919. This result indicates that the estimated risk-testosterone correlation was not significantly different from zero (i.e., null effect). The results are represented by a forest plot in Fig. 2

The $Q_{homogeneity}$ test showed a statistically significant result, highlighting the very heterogeneous set of outcomes: $Q_{homogeneity}$ (93) = 498.6, p < .001. This result demonstrates the necessity for a random effects model, since it allows explaining larger variance compared to a fixed effects model.

3.2. Trim and fill method

The trim and fill method revealed that no measurements (estimate = 0, standard error = 5.71) were trimmed and filled to create a symmetrical funnel plot (Fig. 3). The fail-safe method showed a fail-safe number (Nfs) of 0, i.e. 0 studies with null result should be added in literature to reduce the considered effect size to "non-significantly different from zero". This result is clear since the meta-analysis already estimated a null result.

3.3. Meta-regression - moderators

The meta-regression was run to investigate the influence of any potential moderators. It revealed the statistically significant omnibus effect among the levels of one moderator:

1. Testosterone measure category: Likelihood Ratio (3) = 14.98, p = .002 (Fig. 4)
2. Risk measure category: Likelihood Ratio (6) = 1.992, p = .920 (Fig. 5)

This result represents the significant variability of effect size estimates across different levels of the testosterone measure moderator, when controlling for the risk measure moderator. Nevertheless, when testing all single contrasts, the only statistically significant contrast was represented by the Lottery task level: r = 0.0934, 95 CI [0.0218, 0.1650], z = 2.5564, p = .011. This result shows that the risk-testosterone correlation is significantly higher than the mean value when measured by lottery tasks.

3.4. Meta-regression - sex differences

The meta-regression run on the subset exploring sex differences showed no statistically significant effect of moderation: Likelihood Ratio (1) = 0.1903, p = .663 (Fig. 6), i.e., the risk-testosterone correlation estimate showed no significant differences between male and female participants.

4. Discussion

The present study indicates that testosterone is unlikely to provide a general explanation to human risk preferences. Rather than a single-hormone account, the evidence points to a biopsychosocial view: what we label "risk taking" emerges from the interaction of task demands, cognitive and affective processes, and situational context, with endocrine influences (when detectable) appearing narrow and contingent rather than broad and trait-like. That reading helps reconcile why single studies pointed in different directions, reporting positive, negative, and null links between testosterone and risky choices (Apicella et al., 2008; Stanton et al., 2021; Boksem et al., 2013b)

A central theoretical test in our analysis concerns sex moderation. Although mean differences between males and females are well documented, our meta-analysis finds no evidence that the testosterone-risk link differs by sex. This result challenges sex-specific hormonal accounts: average sex gaps in risk are not, on current evidence, explained by the testosterone-risk coupling. More plausible is a distributed explanation in which socialization, opportunity structures, experience, and cognitive–affective styles account for observed gaps, with hormones modulating behavior only under specific task or situational

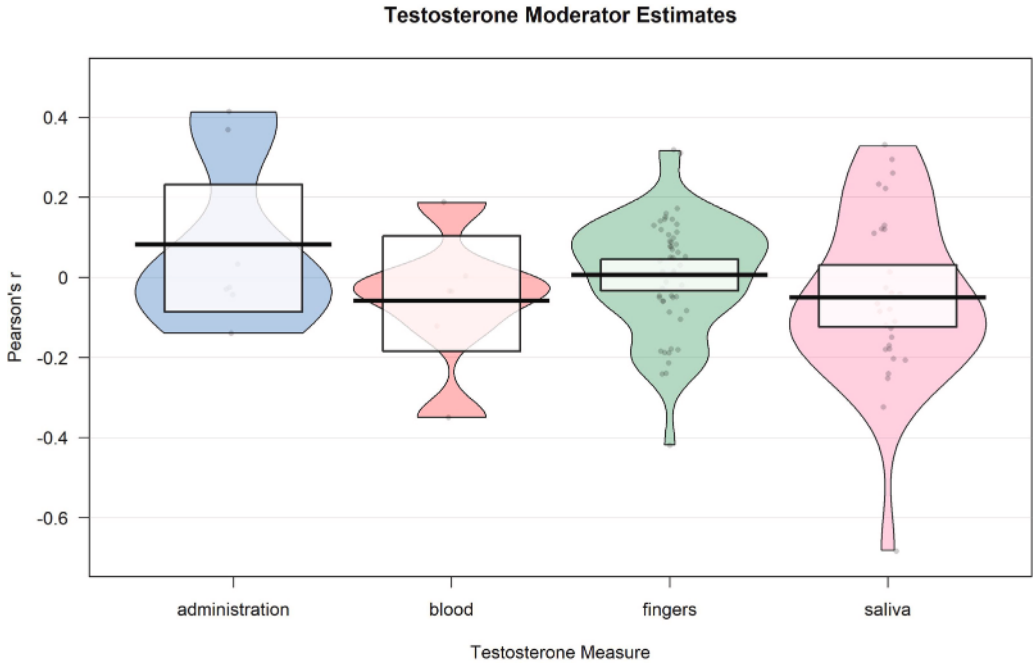

Fig. 4. This plot represents the distribution of effect sizes across different levels of testosterone measure. In this and the following plots, dots represent the raw data (jittered horizontally), the horizontal bar shows central tendencies (i.e., median), beans represent smoothed density distributions, and rectangles represent the 95 confidence intervals.

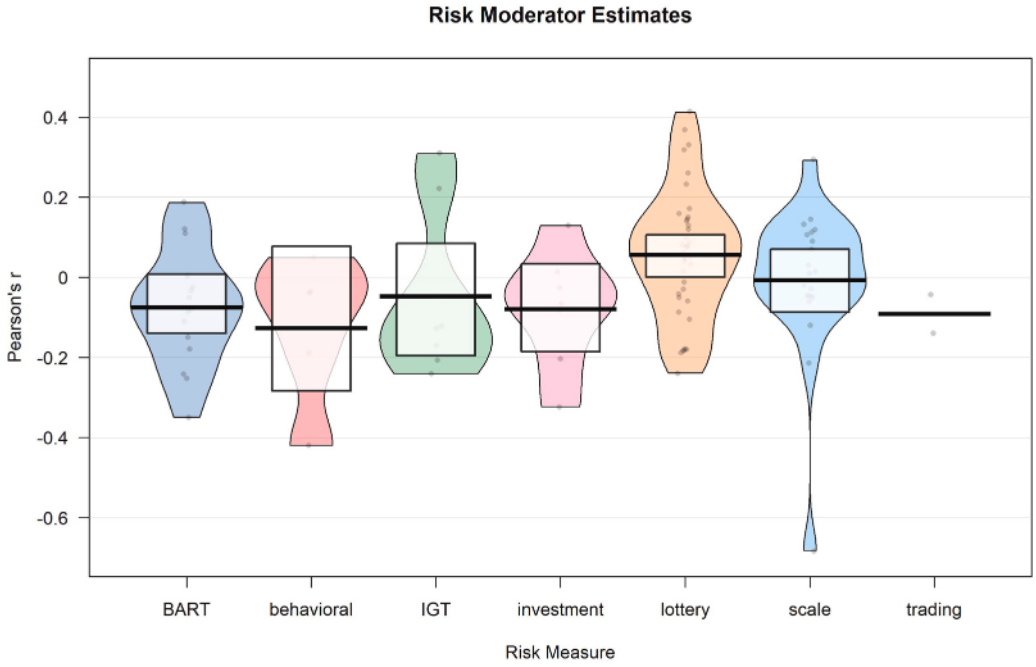

Fig. 5. This plot represents the distribution of effect sizes across different levels of risk measure.

constraints (Byrnes et al., 1999; Eckel and Grossman, 2008). Our results align with this.

Placed against prominent accounts linking androgens to approach motivation, reduced fear, and status-seeking, the present null hypothesis argues against a generalizable route from testosterone to risky choice in economic tasks. Studies in neuroeconomics and behavioral endocrinology have long proposed that testosterone may bias approach and status-oriented behavior (Mazur and Booth, 1998; Stanton, 2017; Apicella et al., 2015), while task-focused work emphasizes that what is labeled "risk" depends on paradigm and affective load (Loewenstein et al., 2001). In light of this framework, the mixed single-study associations appear context (or method) bound rather than constituting a replicable law of behavior across paradigms. Our own results support that interpretation: the grand estimate is null, heterogeneity

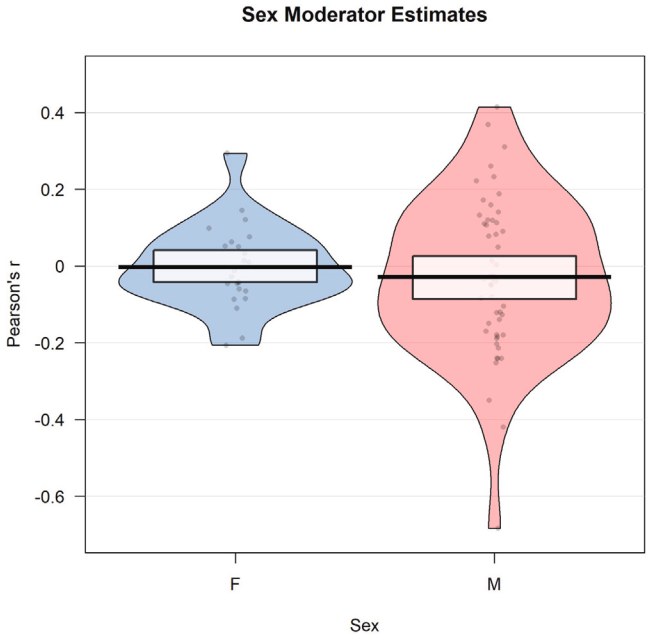

Fig. 6. This plot represents the distribution of effect sizes between male and female participants.

is substantial, and moderator patterns align more with differences in operationalization (task family; endocrine measure) than with a domain-general hormonal mechanism.

The conclusions of our meta-analysis converge with emerging large-scale experimental evidence. In particular, Dreber et al. (2025) conducted a preregistered, double-blind randomized controlled trial in which a large sample of men received either intranasal testosterone or placebo before completing a battery of economic tasks indexing risk preferences, social preferences, and competitiveness. The authors found no evidence that testosterone administration influenced their primary economic preference outcomes, and no strong association between basal salivary testosterone and these preferences within men. Because this trial was published after our pre-specified search end date, it is not included in the quantitative synthesis. Nonetheless, its well-powered null effects mirror the near-zero pooled association we observe here and further undermine the idea that short-term testosterone fluctuations are a general causal driver of economic risk preferences.

Another randomized trial, conducted in a different context, also shows a similar pattern of results. Lieberman et al. (2024) examined the behavioral effects of repeated intramuscular injections of testosterone (200 mg/week) versus placebo during a 28-day period of severe energy restriction in healthy young men. Despite substantial manipulation of circulating testosterone, the authors did not observe consistent effects of exogenous testosterone on a battery of behavioral outcomes, including risk-taking measures, aggression, competition, and cognitive performance. Because this study was also published after our search end date, it is not included in our quantitative synthesis. Nonetheless, its overall pattern of null behavioral findings is aligned with the meta-analytic picture reported here.

Also, "risk" is not a single construct. Classic lottery tasks, sequential learning paradigms like the Iowa Gambling Task, and affect-heavy measures such as the BART each lean on different cognitive and affective processes (Lejuez et al., 2002; Lauriola et al., 2014; Loewenstein et al., 2001). Treating these task families as interchangeable would imply swapping the construct while keeping the label. Consistent with this concern, our moderators indicated that patterns varied by measurement family: lottery-style tasks aligned more with a pro-association narrative, whereas BART, IGT, or self-report measures did not follow suit. This pattern is exactly what we would expect if different paradigms place different weights on expected-value calculation, feedback sensitivity, and arousal (Holt and Laury, 2002; Lauriola et al., 2014; Loewenstein et al., 2001).

On the other hand, endocrine measurement matters just as much. Findings based on direct assays or controlled administration do not align with results derived from indirect proxies such as 2D:4D, whose validity as markers of developmental androgens has been questioned in large, well-powered work (Alonso et al., 2018; Neyse et al., 2021). Where indirect markers hint at associations, direct measures are notably less supportive, a profile that seems more consistent with measurement artifact than with a robust biological pathway (Alonso et al., 2018; Neyse et al., 2021; Branas-Garza and Rustichini, 2011). To test the robustness of these patterns, we conducted a sensitivity analysis restricted to direct measures only (saliva, blood, and experimental testosterone administration), repeating the main meta-analysis and meta-regressions on this subset. As reported in the Supplementary Materials, the overall effect and the key moderator patterns were essentially unchanged compared with the analyses on the complete dataset.

Taken together, these findings support a shift from single-cause biological explanations toward multiplicity. Classic reviews of sex differences in risk behavior emphasize the roles of social learning, incentives, domain specificity, and context sensitivity (Byrnes et al., 1999; Eckel and Grossman, 2008). Our results add that, even when testosterone differs between groups, its covariation with risk behavior does not provide incremental explanatory power at the behavioral level surveyed here. This does not preclude hormonal modulation in narrow states or paradigms; rather, it argues against a trait-like explanatory role for testosterone in the mean differences that motivate much of this literature.

Two design principles follow directly from our moderators. First, *task–mechanism alignment*: specify, a priori, which computational or affective ingredient of "risk" a paradigm targets—expected value computation in Holt–Laury tasks (Holt and Laury, 2002), feedback-driven learning in IGT, or affective arousal in BART (Lauriola et al., 2014; Loewenstein et al., 2001), and analyze endpoints accordingly rather than pooling unlike outcomes. Second, *biological validity*: favor direct endocrine indices or tightly controlled pharmacological manipulations; reserve morphology proxies for questions they can validly address, and report assay metadata, timing, and preprocessing to enable cumulative synthesis; we found out that basic assay details such as sensitivity, intra- and inter-assay coefficients of variation, and sampling time of day were reported inconsistently or were often missing, which limits our ability to run more fine-grained moderator analyses (Alonso et al., 2018; Neyse et al., 2021).

A common worry is that the published literature might hide null results or overrepresent "exciting" findings. If that were true here, a meta-analysis could mistakenly suggest no effect, or miss a small real one. In our case, standard publication bias checks did not show a funnel-plot asymmetry, and our additional evidential diagnostics (i.e., trim and fill method) were consistent with a record that includes its null. In other words, the pattern we see is unlikely to be created by missing studies. The cautious takeaway is that the overall effect really does look absent; where we do see pockets of apparent "signal", they are more plausibly consequences of specific designs and measurements than of a general testosterone pathway (Duval and Tweedie, 2000a,b).

As with any meta-analysis, our conclusions are bounded by the available studies. Differences in sampling schedules, diurnal control, and assay procedures can blur true moment-to-moment links between hormone levels and behavior. Most work in our synthesis relies on a single baseline hormone assessment, typically obtained prior to task performance, which makes it difficult to characterize dynamic endocrine responses to specific task events. By contrast, studies that collect multiple samples across competitive or decision-making episodes

and examine how changes in testosterone (ΔT) relate to behavior (e.g., Mehta and Josephs (2006), Carré et al. (2009), Apicella et al. (2014), Casto et al. (2020), Alacreu-Crespo et al. (2019)) illustrate how phasic endocrine responses can be linked to competitive effort, aggression, or financial risk-taking. Coupling such designs with computational models that decompose risky choice into parameters such as risk sensitivity, loss aversion, and learning rates would allow tests of whether endocrine shifts map onto specific components of decision making rather than global 'risk-taking' scores (e.g., Margittai et al. (2018), Votinov et al. (2022a), Molins et al. (2021). Even though this direction is consistent with the scope of this work, at present this literature is too small and heterogeneous to support a dedicated meta-analysis, but our results highlight it as a key avenue for future research.

As the majority of studies in the literature investigate the impact of testosterone on risk by relating baseline (tonic) testosterone levels, or the effects of an acute single-dose administration, to risk aversion, our analysis primarily reflects this evidence base. A narrower parallel literature examines changes in testosterone (ΔT) across competitive or status-relevant tasks and suggests that phasic endocrine responses may, in some cases, predict risky or competitive behavior more strongly than baseline levels (Mehta and Josephs, 2006; Carré et al., 2009; Apicella et al., 2014). However, these ΔT studies are currently too few and too heterogeneous in design to support a robust meta-analytic synthesis. Quantitatively, our dataset included only one effect size based on ΔT, with no evidence of a strong association with risk-taking. We therefore focus quantitatively on tonic levels and standard single-dose administration protocols, returning to ΔT as an important target for future work. On the behavioral side, "risk" is implemented in multiple ways: varying incentives, framing, and feedback, even within the same family of tasks, and these differences are known to shift behavior (Holt and Laury, 2002; Lauriola et al., 2014). Moreover, the evidence base is concentrated in a limited set of laboratory-like economic paradigms and populations, which can limit generalizability. These constraints, reflected in the heterogeneity summarized in our results, underscore the need for preregistered, multi-measure designs that harmonize endocrine assays and task protocols across laboratories.

Evidence for interactionist (i.e., dual-hormone) accounts remains suggestive but too sparse and uneven to support firm conclusions. The original proposal, that testosterone's effects depend on concurrent cortisol levels (Carre and McCormick, 2008), has motivated several tests, including work showing cortisol-related shifts in financial risk (Kandasamy et al., 2014) and small studies reporting testosterone ×cortisol patterns on laboratory risk tasks or reports (Mehta and Prasad, 2015; Singh, 2021; Ronay et al., 2018). Yet designs, assays (saliva vs. hair), sampling schedules, and task endpoints vary considerably, and our own search did not yield enough comparable studies to meta-analyze the interaction (see Methods). The prudent conclusion is insufficient cumulative evidence at present. Progress will require adequately powered studies that (i) preregister the interaction a priori, (ii) synchronize testosterone and cortisol sampling with task timing, (iii) standardize endocrine assays and report full metadata, and (iv) match risk paradigms to specific computational/affective targets. With these commitments, the dual-hormone hypothesis becomes an empirically tractable program rather than a post hoc narrative.

The most defensible position at present is modest: testosterone is not a universal driver of risk preference. Where associations surface, they are best understood as products of specific constellations of task demands and endocrine measurement rather than a single hormonal lever on decision making. Research programs that embrace multiplicity (i.e., neural computation, learning history, affective state, social context, and carefully specified endocrine pathways) offer a more credible route forward than continued searches for a unitary biochemical signature (Apicella et al., 2008; Stanton et al., 2011b; Zethraeus et al., 2009a).

Financial disclosure

The authors acknowledge financial support from Intesa Sanpaolo Innovation Center S.p.A.

Acknowledgment

The research was conducted under a cooperative agreement between the IMT School for Advanced Studies Lucca and Intesa Sanpaolo Innovation Center S.p.A. and Intesa Sanpaolo S.p.A. The authors would like to thank Giovanna Paladino, Dario Russignaga for Intesa Sanpaolo S.p.A., Sonia D'Arcangelo and Francesca Maggi for Intesa Sanpaolo Innovation Center, for their precious contribution during each project step.

Appendix A. Supplementary Table A.1.

See Table A.1.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.neubiorev.2026.106575.

Table A.1
Database-specific search strategies.

Database	Search query and filters
PubMed	(testosterone[Title/Abstract]) AND (''risk attitude [Title/Abstract] OR ''risk seeking [Title/Abstract] OR ''risk aversion [Title/Abstract]). *Filters:* Humans; English, Spanish, or Italian; no year restrictions. Search fields: Title/Abstract. Search date: 2 March 2023.
Scopus	TITLE-ABS-KEY(testosterone AND (''risk attitude OR ''risk seeking OR ''risk aversion)). *Limits:* Language = English, Spanish, or Italian. Document type: articles and reviews retrieved, with reviews excluded at screening. Search date: 2 March 2023.
Google Scholar	testosterone ''risk attitude OR ''risk seeking OR ''risk aversion . *Settings:* Search performed without date restriction on 2 March 2023; the first 1000 records (titles) were screened.

References

Alacreu-Crespo, A., Costa, R., Abad-Tortosa, D., Hidalgo, V., Salvador, A., Serrano, M.Á., 2019. Hormonal changes after competition predict sex-differentiated decision-making. J. Behav. Decis. Mak. 32 (5), 550–563.

Alonso, J., Di Paolo, R., Ponti, G., Sartarelli, M., 2018. Facts and misconceptions about 2D: 4D, social and risk preferences. Front. Behav. Neurosci. 12, 22.

Apicella, C.L., Carré, J.M., Dreber, A., 2015. Testosterone and economic risk taking: A review. Adapt. Hum. Behav. Physiol. 1, 358–385.

Apicella, C.L., Dreber, A., Campbell, B., Gray, P.B., Hoffman, M., Little, A.C., 2008. Testosterone and financial risk preferences. Evol. Hum. Behav. 29 (6), 384–390.

Apicella, C.L., Dreber, A., Mollerstrom, J., 2014. Salivary testosterone change following monetary wins and losses predicts future financial risk-taking. Psychoneuroendocrinology 39, 58–64.

Boksem, M.A.S., Kostermans, E., Tops, M., De Cremer, D., 2013b. Testosterone inhibits trust but promotes reciprocity. Psychol. Sci. 24 (11), 2306–2314.

Branas-Garza, P., Rustichini, A., 2011. Organizing effects of testosterone and economic behavior: Not just risk taking. PLoS One 6 (12), e29842.

Byrnes, J.P., Miller, D.C., Schafer, W.D., 1999. Gender differences in risk taking: A meta-analysis. Psychol. Bull. 125 (3), 367–383.

Carre, J.M., McCormick, C.M., 2008. The dual-hormone hypothesis of social status and risk taking. Psychol. Sci. 19 (3), 304–308.

Carré, J.M., Putnam, S.K., McCormick, C.M., 2009. Testosterone responses to competition predict future aggressive behaviour at a cost to reward in men. Psychoneuroendocrinology 34 (4), 561–570.

Casto, K.V., Edwards, D.A., Akinola, M., Davis, C., Mehta, P.H., 2020. Testosterone reactivity to competition and competitive endurance in men and women. Horm. Behav. 123, 104665.

Cohen, J., Cohen, P., West, S.G., Aiken, L.S., 2013. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Routledge, New York, NY.

Del Re, A.C., 2013. compute.es: Compute effect sizes. R package version 0.2-2. URL https://cran.r-project.org/package=compute.es.

Derntl, B., Windischberger, C., Robinson, S., Kryspin-Exner, I., Gur, R.C., Moser, E., 2014b. Effects of sex steroid hormones on decision-making in women and men: Role of dopaminergic neuromodulation. Front. Neurosci. 8, 452.

Dreber, A., Johannesson, M., Nave, G., Apicella, C., Geniole, S., Imai, T., Knight, E., Manfredi, D., Mehta, P., Proietti, V., Stanton, S., Zeltikova, A., Luberti, F., Ortiz, T., Carré, J., 2025. Investigating the effects of single-dose intranasal testosterone on economic preferences in a large randomized trial of men. Proceed. National Academy Sci. 122 (39), e2508519122.

Duval, S., Tweedie, R., 2000a. A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. J. Amer. Statist. Assoc. 95 (449), 89–98.

Duval, S., Tweedie, R., 2000b. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56 (2), 455–463.

Eckel, C.C., Grossman, P.J., 2008. In: Plott, C.R., Smith, V.L. (Eds.), Chapter 113 Men, Women and Risk Aversion: Experimental Evidence. In: Handbook of Experimental Economics Results, vol. 1, Elsevier, pp. 1061–1073, URL https://www.sciencedirect.com/science/article/pii/S1574072207001138.

Friedl, A., Pondorfer, A., Schmidt, U., 2020. Gender differences in social risk taking. J. Econ. Psychol. 77, 102182.

Guiso, L., Paiella, M., 2008. Risk aversion, wealth, and background risk. J. Eur. Econ. Assoc. 6 (6), 1109–1150.

Heidari, S., Babor, T.F., De Castro, P., Tort, S., Curno, M., 2016. Sex and gender equity in research: Rationale for the SAGER guidelines and recommended use. Res. Integr. Peer Rev. 1 (1), 2.

Holt, C.A., Laury, S.K., 2002. Risk aversion and incentive effects. Am. Econ. Rev. 92 (5), 1644–1655.

Jackson, C.J., et al., 2005. Personality and risk taking: The role of the Big Five personality traits. J. Pers. Soc. Psychol. 89 (6), 1313–1326.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47 (2), 263–291.

Kandasamy, N., Hardy, B., Page, L., Schaffner, M., Graggaber, J., Sailer, U., et al., 2014. Cortisol shifts financial risk preferences. Proc. Natl. Acad. Sci. 111 (9), 3608–3613.

Lakdawalla, D., Philipson, T., Jena, A., 2020. Health technology assessment with risk aversion in health. Health Econ. 29 (S1), 4–15.

Lauriola, M., Panno, A., Levin, I.P., Lejuez, C.W., 2014. Individual differences in risky decision-making: A meta-analysis of sensation seeking and impulsivity with the balloon analogue risk task. J. Behav. Decis. Mak. 27 (1), 20–36.

Lejuez, C.W., Read, J.P., Kahler, C.W., Richards, J.B., Ramsey, S.E., Stuart, G.L., Strong, D.R., Brown, R.A., 2002. Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). J. Exp. Psychol.: Appl. 8 (2), 75.

Lieberman, H.R., Caldwell, J.A., Vartanian, O., Carmichael, O.T., Karl, J.P., Berryman, C.E., Gadde, K.M., Niro, P.J., Harris, M.N., Rood, J.C., Pasiakos, S.M., 2024. Effects of testosterone enanthate on aggression, risk-taking, competition, mood, and other cognitive domains during 28 days of severe energy deprivation. Psychopharmacology 241 (3), 461–478, Epub 2023 Dec 1.

Loewenstein, G.F., Weber, E.U., Hsee, C.K., Welch, N., 2001. Risk as feelings. Psychol. Bull. 127 (2), 267–286.

Margittai, Z., Nave, G., Van Wingerden, M., Schnitzler, A., Schwabe, L., Kalenscher, T., 2018. Combined effects of glucocorticoid and noradrenergic activity on loss aversion. Neuropsychopharmacology 43 (2), 334–341.

Mazur, A., Booth, A., 1998. Testosterone and dominance in men. Behav. Brain Sci. 21 (3), 353–397.

Mehta, P.H., Josephs, R.A., 2006. Testosterone change after losing predicts the decision to compete again. Horm. Behav. 50 (5), 684–692.

Mehta, P.H., Prasad, S., 2015. The dual-hormone hypothesis: A brief review and future research agenda. Curr. Opin. Behav. Sci. 3, 163–168.

Molins, F., Ayuso, C., Serrano, M.Á., 2021. Early stages of the acute physical stress response increase loss aversion and learning on decision making: A Bayesian approach. Physiol. Behav. 237, 113459.

Mueller, D., et al., 2006. The role of extraversion and neuroticism in risky decision making. Pers. Individ. Differ. 40 (5), 841–853.

Mullen, B., 1989. Advanced Basic Meta-Analysis: Version 1.10. Lawrence Erlbaum Associates, Hillsdale, NJ.

Neyse, L., Johannesson, M., Dreber, A., 2021. 2D: 4D does not predict economic preferences: Evidence from a large, representative sample. J. Econ. Behav. Organ. 185, 390–401.

Rai, J., Kimmel, J., 2015. Gender differences in risk preferences: An empirical study using attitudinal and behavioral specifications of risk aversion. In: Gender in the Labor Market. Emerald Group Publishing Limited, pp. 61–91.

Ramirez, M., et al., 2003. Predicting risk taking behavior from the five factor model of personality. Pers. Individ. Differ. 36 (3), 695–705.

Ronay, R., Van der Meij, L., Oostrom, J.K., Pollet, T.V., 2018. No evidence for a relationship between hair testosterone concentrations and 2D: 4D ratio or risk taking. Front. Behav. Neurosci. 12, 30.

Rosenthal, R., 1991. Meta-Analytic Procedures for Social Research. SAGE Publications, Inc., Thousand Oaks, CA.

RStudio Team, 2023. RStudio: Integrated development environment for R. Version 2023.06.1. URL https://posit.co.

Sapienza, P., Zingales, L., Maestripieri, D., 2009a. Gender differences in financial risk aversion and career choices are affected by testosterone. Proc. Natl. Acad. Sci. 106 (36), 15268–15273.

Sarin, R., Wieland, A., 2016. Risk aversion for decisions under uncertainty: Are there gender differences? J. Behav. Exp. Econ. 60, 1–8, URL https://www.sciencedirect.com/science/article/pii/S2214804315001305.

Singh, V., 2021. Role of cortisol and testosterone in risky decision-making: deciphering male decision-making in the iowa gambling task. Front. Neurosci. 15, 631195.

Soane, E., Chmiel, N., 2005. Are risk preferences consistent?: The influence of decision domain and personality. Pers. Individ. Differ. 38 (8), 1781–1791.

Stanton, S.J., 2017. The role of testosterone and estrogen in consumer behavior and social & economic decision making: a review. Horm. Behav. 92, 155–163.

Stanton, S.J., Liening, S.H., Schultheiss, O.C., 2011a. Testosterone is positively associated with risk taking in the Iowa Gambling Task. Horm. Behav. 59 (2), 252–256.

Stanton, S.J., Mullette-Gillman, O.A., McLaurin, R.E., Kuhn, C.M., LaBar, K.S., Platt, M.L., Huettel, S.A., 2011b. Low- and high-testosterone individuals exhibit decreased aversion to economic risk. Psychol. Sci. 22 (4), 447–453.

Stanton, S.J., Welker, K.M., Bonin, P.L., Goldfarb, B., Carré, J.M., 2021. The effect of testosterone on economic risk-taking: A multi-study, multi-method investigation. Horm. Behav. 134, 105014.

Stenstrom, E.P., Dinsmore, J.B., Kunstman, J.W., Vohs, K.D., 2018. The effects of money exposure on testosterone and risk-taking, and the moderating role of narcissism. Pers. Individ. Differ. 123, 110–114.

Stenstrom, E., Saad, G., Nepomuceno, M.V., Mendenhall, Z., 2011. Testosterone and domain-specific risk: Digit ratios (2D:4D and rel2) as predictors of recreational, financial, and social risk-taking behaviors. Pers. Individ. Differ. 51 (4), 412–416.

Strobel, M., et al., 2001. Psychometric properties of the risk taking inventory. Psychol. Assess. 8 (2), 75–85.

Tversky, A., Kahneman, D., 1994. Decision making under risk. J. Exp. Psychol. [Gen.] 123 (4), 507–519.

Viechtbauer, W., 2010. Conducting meta-analyses in r with the metafor package. J. Stat. Softw. 36 (3), 1–48.

Von Neumann, J., Morgenstern, O., 1944. Theory of Games and Economic Behavior, first ed. Princeton University Press, Princeton, NJ.

Votinov, M., Knyazeva, I., Habel, U., Konrad, K., Puiu, A.A., 2022a. A Bayesian modeling approach to examine the role of testosterone administration on the endowment effect and risk-taking. Front. Neurosci. 16, 858168.

Zethraeus, N., Kocoska-Maras, L., Ellingsen, T., von Schoultz, B., Hirschberg, A.v.L., Johannesson, M., 2009a. A randomized trial of the effect of estrogen and testosterone on economic behavior. Proc. Natl. Acad. Sci. 106 (16), 6535–6538.

Additional articles included in the meta analysis

Anderson, T., 2012. Comparing risk-taking and digit ratio (2D: 4D) in offenders and non-offenders. Plymouth Stud. Sci. 5 (2), 105–120.

Anderson, L.R., Mellor, J.M., 2008. Predicting health behaviors with an experimental measure of risk preference. J. Health Econ. 27 (5), 1260–1274.

Aycinena, D., Baltaduonis, R., Rentschler, L., 2014. Risk preferences and prenatal exposure to sex hormones for ladinos. PLoS One 9 (8), e103332.

Bessey, D., 2021. Loss aversion and health behaviors: Results from two incentivized economic experiments. Healthcare 9 (8), 1040.

Brañas-Garza, P., Galizzi, M.M., Nieboer, J., 2014. Digit ratio and risk taking: evidence from a large, multi-ethnic sample.

Brañas-Garza, P., Galizzi, M.M., Nieboer, J., 2018. Experimental and self-reported measures of risk taking and digit ratio (2d: 4d): evidence from a large, systematic study. Internat. Econom. Rev. 59 (3), 1131–1157.

Candelo, N., Eckel, C., 2018. The 2D: 4D ratio does not always correlate with economic behavior: A field experiment with african-Americans. Econ. Hum. Biology 30, 172–181.

Chicaiza-Becerra, L.A., Garcia-Molina, M., 2017. Prenatal testosterone predicts financial risk taking: Evidence from Latin America. Pers. Individ. Differ. 116, 32–37.

Cueva, C., Roberts, R.E., Spencer, T., Rani, N., Tempest, M., Tobler, P.N., et al., 2015. Cortisol and testosterone increase financial risk taking and may destabilize markets. Sci. Rep. 5 (1), 11206.

Derntl, B., Pintzinger, N., Kryspin-Exner, I., Schöpf, V., 2014a. The impact of sex hormone concentrations on decision-making in females and males. Front. Neurosci. 8, 352.

Desmoulins-Lebeault, F., Gajewski, J.F., Meunier, L., 2019. Do incentives contracts lead to higher risk-taking? The impact of executives' characteristics. SSRN.

Ellis, L., Hoskin, A.W., Ratnasingam, M., 2016. Testosterone, risk taking, and religiosity: Evidence from two cultures. J. Sci. Study Relig. 55 (1), 153–173.

Evans, K.L., Hampson, E., 2014. Does risk-taking mediate the relationship between testosterone and decision-making on the iowa gambling task? Pers. Individ. Differ. 61, 57–62.

Garbarino, E., Slonim, R., Sydnor, J., 2010. Digit ratios predict risk aversion for both sexes. ACR North Am. Adv

Garbarino, E., Slonim, R., Sydnor, J., 2011. Digit ratios (2D: 4D) as predictors of risky decision making for both sexes. J. Risk Uncertain. 42, 1–26.

Goudriaan, A.E., Lapauw, B., Ruige, J., Feyen, E., Kaufman, J.M., Brand, M., Vingerhoets, G., 2010. The influence of high-normal testosterone levels on risk-taking in healthy males in a 1-week letrozole administration study. Psychoneuroendocrinology 35 (9), 1416–1421.

Kastlunger, B., Dressler, S.G., Kirchler, E., Mittone, L., Voracek, M., 2010. Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2D: 4D). J. Econ. Psychol. 31 (4), 542–552.

Kim, Y., Kim, K., Kim, T.H., 2014. Domain specific relationships of 2D: 4D digit ratio in risk perception and risk behavior. J. Gen. Psychol. 141 (4), 373–392.

Lam, D., Ozorio, B., 2015. An exploratory study of the relationship between digit ratio, illusion of control, and risk-taking behavior among Chinese college students. J. Gambl. Stud. 31, 1377–1385.

Levy, I., Tokoglu, F., Kritzer, M.F., Frank, M.J., Schonberg, T., Daw, N.D., et al., 2010. Neural representation of subjective value under risk and ambiguity. J. Neurophysiol. 103 (2), 1036–1047.

Lima de Miranda, K., Neyse, L., Schmidt, U., 2018. Risk preferences and predictions about others: No association with 2D: 4D ratio. Front. Behav. Neurosci. 12, 9.

Mehta, P.H., Welker, K.M., Zilioli, S., Carré, J.M., 2015. Testosterone and cortisol jointly modulate risk-taking. Psychoneuroendocrinology 56, 88–99.

Millet, K., Dewitte, S., Vantomme, D., 2005. Digit Extension: Validation of a New Biometric Variable. TEW Research Report 0546.

Nofsinger, J.R., Patterson, F.M., Shank, C.A., 2018. Decision-making, financial risk aversion, and behavioral biases: The role of testosterone and stress. Econ. Hum. Biol. 29, 1–16.

Nofsinger, J.R., Patterson, F.M., Shank, C.A., 2021. On the physiology of investment biases: the role of cortisol and testosterone. J. Behav. Financ. 22 (3), 338–349.

Open Science Collaboration, 2015. Estimating the reproducibility of psychological science. Science 349 (6251), aac4716.

Oran, J.S., Akyatan, A., 2012. A pilot study for measuring correlations between hormone levels and risk taking in men and women at different times of day. Int. J. Behav. Account. Financ. 202–220.

Patterson, F.M., 2014. The relation of steroid hormones and personality factors to financial performance and risk-taking behavior.

Peper, J.S., Koolschijn, P.C.M., Crone, E.A., 2013. Development of risk taking: contributions from adolescent testosterone and the orbito-frontal cortex. J. Cogn. Neurosci. 25 (12), 2141–2150.

Ronay, R., Von Hippel, W., 2010. Power, testosterone, and risk-taking. J. Econom. Theory 23 (5), 473–482.

Rothschild, M., Stiglitz, J., 1970. Increasing risk: I. A definition. J. Econom. Theory 2 (3), 225–243.

Sapienza, P., Zingales, L., Maestripieri, D., 2009b. Gender differences in financial risk aversion and career choices are affected by testosterone. Proc. Natl. Acad. Sci. 106 (36), 15268–15273.

Schipper, B.C., 2012. Sex hormones and choice under risk (no. 12–7). In: Working Paper.

Smith, K.M., Apicella, C.L., 2017. Winners, losers, and posers: The effect of power poses on testosterone and risk-taking following competition. Horm. Behav. 92, 172–181.

Thaler, R.H., 1980. Toward a positive theory of consumer choice. J. Econ. Behav. Organ. 1 (1), 39–60.

Topel, S., Kortink, E.D., Liu, H., Cavanagh, J.F., van der Molen, M.J.W., 2024. Frontal-midline theta promotes context-dependent risk aversion in social anxiety.

Trahms, C.A., Coombs, J.E., Barrick, M., 2010. Does biology matter? How prenatal testosterone, entrepreneur risk propensity, and entrepreneur risk perceptions influence venture performance. Front. Entrep. Res. 30 (5), 4.

Vermeersch, H., T'sjoen, G., Kaufman, J.M., Vincke, J., 2008. The role of testosterone in aggressive and non-aggressive risk-taking in adolescent boys. Horm. Behav. 53 (3), 463–471.

Votinov, M., Radke, S., Heimann, K., Krämer, B., et al., 2022b. Testosterone modulates economic risk-taking: Evidence from a pre-registered multi-study investigation. Front. Neurosci. 16, 858168.

Wagels, L., Votinov, M., Radke, S., Clemens, B., Montag, C., Jung, S., Habel, U., 2017. Blunted insula activation reflects increased risk and reward seeking as an interaction of testosterone administration and the MAOA polymorphism. Hum. Brain Mapp. 38 (9), 4574–4593.

Welker, K.M., Roy, A.R., Geniole, S., Kitayama, S., Carré, J.M., 2019. Taking risks for personal gain: An investigation of self-construal and testosterone responses to competition. Soc. Neurosci. 14 (1), 99–113.

Wu, Y., Clark, L., Zilioli, S., Eisenegger, C., Gillan, C.M., Deng, H., Li, H., 2018. Single dose testosterone administration modulates emotional reactivity and counterfactual choice in healthy males. Psychoneuroendocrinology 90, 127–133.

Xie, Z., Page, L., Hardy, B., 2017. Investigating gender differences under time pressure in financial risk taking. Front. Behav. Neurosci. 11, 246.

Zethraeus, N., Kocoska-Maras, L., Ellingsen, T., Von Schoultz, B.O., Hirschberg, A.L., Johannesson, M., 2009. A randomized trial of the effect of estrogen and testosterone on economic behavior. Proc. Natl. Acad. Sci. 106 (16), 6535–6538.

Kim, Y., Kim, K., Kim, T.H., 2014. Domain specific relationships of 2D:4D digit ratio in risk perception and risk behavior. The Journal of General Psychology 141 (4), 373–392.

Welker, K.M., Roy, A.R.K., Geniole, S., Kitayama, S., Carré, J.M., 2019. Taking risks for personal gain: An investigation of self-construal and testosterone responses to competition. Social Neuroscience 14 (1), 99–113.