1	A meta-analysis of the association between male dimorphism and fitness
2	outcomes in humans
3	Lidborg, L. H.1*, Cross, C. P.2, Boothroyd, L. G.1
4 5 6	 Department of Psychology, Durham University, South Road, Durham, DH1 3LE, UK School of Psychology and Neuroscience, University of St Andrews, St Mary's St Mary's Quad, South St, St Andrews KY16 9JP, UK
7	*Author for correspondence: <u>linda.h.lidborg@durham.ac.uk</u>
8 9 10 11	Word count: 8,279 Keywords: sexual selection; human evolution; sexual dimorphism; masculinity; mating success; reproductive success
12 13	Abstract
14	Humans are sexually dimorphic: men and women differ in body build and composition,
15	craniofacial structure, and voice pitch, likely mediated in part by developmental testosterone.
16	Sexual selection hypotheses posit that, ancestrally, more 'masculine' men may have acquired
17	more mates and/or sired more viable offspring. Thus far, however, evidence for either
18	association is unclear. Here, we meta-analyze the relationships between six masculine traits
19	and mating/reproductive outcomes (96 studies, 474 effects, $N = 177,044$). Voice pitch, height,
20	and testosterone all predicted mating; however, strength/muscularity was the strongest and
21	only consistent predictor of both mating and reproduction. Facial masculinity and digit ratios
22	did not significantly predict either. There was no clear evidence for any effects of masculinity
23	on offspring viability. Our findings support arguments that strength/muscularity may be
24	sexually selected in humans, but cast doubt regarding selection for other forms of masculinity
25	and highlight the need to increase tests of evolutionary hypotheses outside of industrialized
26	populations.

27 Introduction

28

29 Sexual dimorphism and masculinity in humans

Sexual dimorphism refers to sex differences in morphological and behavioral traits, excluding 30 reproductive organs (1), with particular emphasis on traits thought to have evolved through 31 sexual selection (2). Humans are a sexually dimorphic species (1). Sexual selection in 32 33 mammalian species, including human and non-human primates, is commonly argued to have acted more strongly on male traits, as a consequence of greater variance in males' 34 reproductive output (3) and a male-biased operational sex ratio, i.e. a surplus of 35 reproductively available males relative to fertile females (e.g. 4). 36 Dimorphic traits that are exaggerated in males are typically referred to as masculine. In 37 humans, masculine faces are characterized by features such as a pronounced brow ridge, a 38 longer lower face, and wider mandibles, cheekbones, and chins (5). Men are, on average, 7-39 8% taller than women (6) and weigh approximately 15% more (7). Relative to this fairly 40 41 modest body size dimorphism, upper body musculature and strength are highly dimorphic in humans: compared to women, men have 61% more overall muscle mass, and 90% greater 42 upper body strength (8). Men's bodies also tend to have a V- or wedge-shape, showing a 43 44 greater shoulder-to-hip ratio (9, 10) and waist-to-chest ratio (11, 12) than women's. Secondto-fourth finger (digit) length ratios are often claimed to be sexually dimorphic, with men's 45 2D:4D typically being lower than women's (13; though this may not be universal: 14). In 46 47 addition, fundamental frequency, commonly referred to as voice pitch, is nearly six standard deviations lower in men than in women (15). 48

The development of these masculine traits in men is influenced by exposure to
androgens, particularly testosterone. With the exception of 2D:4D, which is commonly
claimed to be influenced primarily by prenatal testosterone levels and is present at birth (16;

but see 17), masculine traits generally develop or become exaggerated following a surge in
testosterone production at sexual maturity (18-20) – although it is not necessarily clear
whether the size of that surge corresponds directly to the extent of trait expression.

55 Proposed mechanisms underlying the evolution of masculine traits

Key to the assumption that men's masculine traits are sexually selected is that masculine traits should be reliably associated with greater biological fitness. Men may increase fitness by producing a greater quantity of offspring overall (i.e. greater *fertility*), by acquiring a greater number of partners which may in turn mediate offspring numbers (greater *mating success*), and/or by producing more *surviving* offspring (greater *reproductive success*).

Two key hypotheses and attendant mechanisms have been drawn on by evolutionary 61 behavioral scientists, predicting positive associations between masculinity and fitness 62 outcomes. Firstly, according to the immunocompetence handicap hypothesis (21), masculine 63 traits are a costly signal of heritable immunocompetence, i.e. good genetic quality, due to the 64 65 putative immunosuppressive properties of testosterone (see 22). Masculine men should therefore produce healthier and more viable offspring, who are more likely to survive. Thus, 66 women should be able to increase their fitness (via offspring survival) by selecting masculine 67 68 men as mates. Authors therefore suggested that masculinity in men is intersexually selected, evolved and/or maintained through female choice, and should be associated with greater 69 mating success in contexts where women are able to exercise choice. This should thus result 70 in greater reproductive success, and an advantage in offspring survival. 71

The immunocompetence handicap hypothesis has persisted in the literature, particularly with reference to facial masculinity (although there are no *a priori* reasons to expect this putative mechanism to act more strongly on men's faces than on their bodies), despite concerns regarding its validity since at least 2005 (23). While beyond the scope of this article, common criticisms include that the relationship between testosterone and health is complex
(24), and facial masculinity is inconsistently linked to health (e.g. 25-29). Evidence is
similarly mixed regarding the key assumption that women are attracted to masculinity in
men's faces (25, 30) and bodies (31-34).

Secondly, under the *male-male competition hypothesis*, authors have argued that
formidable (i.e. physically strong and imposing) men are better equipped to compete with
other men for resources, status, and partners (35, 36), through e.g. direct physical contests or
by deterring rivals indirectly (35, 37). For instance, increased musculature may intimidate
competitors by signaling fighting prowess (38) and strength (39), while facial masculinity and
voice pitch may also have an indirect relationship with perceived formidability (40-45).

Importantly, while male-male competition is often framed as an alternative to female 86 choice, women may preferentially mate with both well-resourced men, and with competitive 87 men, facilitating intersexual selection for masculinity (i.e. a 'sexy sons' effect, see 46) where 88 male status is due to, or competitiveness is cued by, formidability (28). Some authors have 89 90 suggested that formidability increases men's mating success through dominance over other 91 men (which may create the circumstances that women select them as mates) rather than women's direct preferences for formidable traits per se (47-49). However, regardless of 92 93 whether the driving mechanism is intra- or intersexual selection (or a combination thereof), the male-male competition hypothesis predicts that formidable men will acquire more partners 94 over their lifetime, which will in turn result in more offspring. This approach, however, does 95 not make any particular predictions regarding offspring health or survival. 96

97 It can be noted that proponents of both the immunocompetence and male-male
98 competition hypotheses have also suggested that more masculine men may show reduced
99 investment in romantic relationships and in offspring (50-53), potentially suppressing
100 offspring health/survival. This could arise from an association between circulating

testosterone (which masculine traits are commonly argued to index) and motivation for sexual 101 102 behavior (54, 55) shifting effort away from parental investment towards pursuit of mating opportunities. Two important caveats here, however, are that the relationship between men's 103 testosterone levels in adolescence (when most masculine traits become exaggerated) and in 104 105 adulthood is exceedingly weak (56), and masculine trait expression in adulthood is not consistently correlated with adult testosterone levels (e.g. 57, 58). Simply being more 106 107 attractive to potential new partners, however, might shift behavior away from relationship investment (for discussion see e.g. 59). Because of this, many authors have previously 108 suggested that women face a trade-off between the (health or competitive) benefits of 109 110 masculinity, and paternal investment.

111 The association between masculine traits and biological fitness

We therefore have at least two theoretical positions which assert that masculine men should 112 have greater numbers of sexual partners, greater offspring numbers, and perhaps a greater 113 114 proportion of surviving offspring, in at least some circumstances. Studies addressing these predictions in societies without effective contraception have done so directly via offspring 115 numbers and/or offspring survival. In most industrialized populations, where access to 116 117 contraceptives attenuates the relationship between sexual behavior and reproductive success, mating success measures are often used instead. These include preferences for casual sex, 118 number of sexual partners, and age at first sexual intercourse (earlier sexual activity allows 119 for a greater lifetime number of sexual partners), as these are assumed to have correlated with 120 reproductive success in men under ancestral conditions (60). 121

A key problem, however, is that the predictions outlined above do not always capture the diversity of human reproductive ecologies even where diverse data exists. We have already noted the fact that female choice may be important to outcomes above. Furthermore, even amongst non-contracepting populations, differences in rates of polygyny, pair-bond

breakdown, and attitudes to fertility may moderate reproductive success and its variance. For 126 127 instance, monogamous cultures do not typically show greater variance in men's versus women's reproductive success (61) and while increasing numbers of sexual partners (e.g. in 128 serially monogamous or polygynous cultures) may often be important for increasing male 129 130 reproductive success, the inverse is true amongst the Pimbwe where women are more advantaged by increased numbers of partners (62). Similarly, although the strongly 131 132 monogamous Agta show high rates of fertility (63), data from ostensibly non-contracepting rural Catholics in C20th Poland (64) shows much lower rates of fertility. These issues 133 highlight the fact that humans have likely had diverse reproductive and pair-bonding norms 134 135 for a long time. As such we can make two observations. Firstly, availability of contraception 136 in low-fertility samples might 'free' sexual behavior from the constraints of pregnancy avoidance, and we might find stronger relationships between any evolved motivation for sex, 137 and actual sexual behavior, in these samples than would have necessarily been found 138 ancestrally. Secondly, however, any adaptation which has been maintained across recent 139 140 hominid lineages must have been adaptive on average across diverse reproductive ecologies. As such, if the proposed adaptation (masculinity leading to enhanced reproductive success via 141 142 mating, and possibly increased offspring survival) exists, we should expect to see both: *i*. 143 masculinity being associated with increased mating success in both high and (perhaps especially) low fertility populations, and *ii*. masculinity being on average positively associated 144 with fertility, and potentially offspring survival, in non-contraception/high fertility 145 146 populations.

147 Meta-analysis in sexual selection

Meta-analysis can be a valuable tool in understanding overall patterns in evolutionarily relevant traits, both across and within species. Jennions and colleagues (65) noted that many traits hypothesized to predict male mating success had not been subject to meta-analysis, and

further argued that while such meta-analyses can be valuable in clarifying the nature and 151 152 extent of selection for some traits, at other times they act to refute prior assumptions. They say: "A general insight from sexual selection meta-analyses ... is that it is easy to be misled 153 by a few high-profile studies into believing that a prediction is well supported. Support is 154 often weaker than assumed." (p.1139). This point does not just apply to comparative research, 155 but is relevant to human sexual selection work specifically. For instance, van Dongen and 156 157 Gangestad (66) found that evidence for health benefits of symmetry were weaker and harder to demonstrate meta-analytically than they would have supposed, given the size of the extant 158 literature. Similarly, when two meta-analyses into the effects of menstrual cycle on women's 159 160 behavior, mate preferences, and attractiveness reached opposing conclusions (67, 68), the 161 exercise suggested that some cycle effects were unlikely to be robust. Indeed, the more cautious analytical methods (e.g. treating unknown null results as zero rather than excluding 162 163 them from analysis) resulted in a null overall effect -a finding that was later borne out by multiple large, pre-registered, studies (69-71). The authors of the meta-analysis that found a 164 null effect suggested that publication and inclusion bias was a particular problem in the field 165 (72), although others argued against this (73). 166

In terms of the current topic, previous studies explicitly testing the relationships 167 between masculine traits and fitness outcomes have been overwhelmingly conducted in low 168 fertility samples and have produced a mixture of positive, negative, and null results (e.g. 63, 169 74, 75). This creates a clear need for meta-analytic comparison of evidence from as wide a 170 population sample as possible. To date, however, meta-analytic analyses are rare, typically 171 172 exclude many aspects of masculinity, and focus on *either* mating or reproductive outcomes, 173 despite both being relevant to testing the theories above. Van Dongen and Sprengers (76) meta-analyzed the relationships between men's handgrip strength (HGS) and sexual behavior 174 in only three industrialized populations (showing a weak, positive association [r = .24]). 175

Across 33 non-industrialized societies, von Rueden and Jaeggi (77) found that male status 176 (which included, but was not limited to, measures of height and strength) weakly predicted 177 reproductive success (overall r = .19). In contrast, Xu and colleagues (78) reported no 178 significant association between men's height and offspring numbers across 16 studies when 179 180 analyzing both industrialized and non-industrialized populations. Lastly, Grebe and colleagues' (54) meta-analysis of 16 effects – the majority of which came from Western 181 samples - showed that men with high levels of circulating testosterone, assayed by blood or 182 saliva, invested more in mating effort, indexed by mating with more partners and showing 183 greater interest in casual sex (r = .22). Across all of their analyses (which also included pair-184 185 bond status, fatherhood status, and fathering behaviors), Grebe and colleagues found no 186 significant differences between 'Western' and 'non-Western' samples, but their 'non-Western' grouping for the relevant analysis only included a low fertility population in 21st 187 Century China. To our knowledge, facial masculinity, voice pitch, and 2D:4D have never 188 been meta-analyzed in relation to mating and/or reproduction. 189

190 The present study

The present article therefore searched widely for published and unpublished data to 191 meta-analyze the relationships between six main masculine traits in men (facial masculinity, 192 body masculinity, 2D:4D, voice pitch, height, and testosterone levels) and both mating and 193 reproductive outcomes, in both high and low fertility samples. By including multiple traits, a 194 broad search strategy, and considering high and low fertility samples both separately and 195 together, we can ascertain whether the current scientific evidence base provides plausible 196 support for the sexual selection of masculine traits in humans. By further testing the 197 publication status of each effect (whether the specific effect size/analysis was reported in a 198 published article or not), we can also evaluate the evidence for publication bias, since this is 199 known to artificially inflate effects in diverse literatures. 200

Mating measures included behavioral measures such as number of sexual partners, 201 202 number of marital spouses, and age at first sexual intercourse. Since increased mating effort is an additional possible route to increased reproductive output, we also included mating 203 204 attitudes, such as preferences for casual sex. Reproductive measures included: fertility measures, such as number of children/grandchildren born and age at the birth of the first 205 child; and reproductive success measures, i.e. number of offspring surviving childhood. Since 206 207 offspring mortality is a measure specifically of offspring viability, we included this as a separate measure (i.e. mortality rate and/or number of deceased offspring). 208

209 Methods

210 Literature search and study selection

A systematic search was initially carried out between November 2017 and February 2018 using the databases PsycINFO, PubMed, and Web of Science; the searches were saved and search alerts ensured inclusion of subsequently published studies. Search terms are given in Box 1.

(masculin* OR "sexual dimorphism" OR "sexually dimorphic" OR width-to-height OR
muscularity OR shoulder-to-hip OR chest-to-waist OR "digit ratio" OR 2d:4d OR "hand grip
strength" OR "handgrip strength" OR "grip strength" OR testosterone OR "voice pitch" OR
"vocal pitch" OR voice OR "non-fat body mass" OR "lean body mass" OR "fundamental
frequency" OR "facial* dominan*" OR height OR "sexual dimorphism in stature" OR "CAG
repeat*")

221 AND

222 ("sex* partner*" OR "short-term relationship*" OR "short term mating" OR "extra pair"
223 OR sociosexual* OR "age of first intercourse" OR "age of first sexual intercourse" OR "age

at first intercourse" OR "age at first sexual intercourse" OR "age of sexual debut" OR "age
at first sex" OR "mating success" OR "number of offspring" OR "offspring number" OR
"number of children" OR "number of grandoffspring" OR "number of grand offspring" OR
"offspring health" OR "offspring mortality" OR "mortality of offspring" OR "surviving
offspring" OR "offspring survival" OR "reproductive onset" OR "reproductive success" OR
"long-term relationship*" OR "age of first birth")

- 230 *AND (human OR man OR men OR participant*).*
- 231 Box 1. Search terms for meta-analysis study discovery

Studies were also retrieved through cross-referencing, citation searches/alerts, and by asking 232 for data on social media. The systematic search generated 2,221 results, including duplicates, 233 and a further approximately 300 articles were found by other means. After scanning titles and 234 abstracts, 280 articles/dissertations were reviewed in full. Studies submitted up to 1 May 2020 235 were accepted. Eligible studies included at least one of the following predictors: facial 236 masculinity, body masculinity (strength, body shape, or muscle mass/non-fat body mass), 237 2D:4D, voice pitch, height, or testosterone levels. The following outcome measures were 238 239 included:

- Mating domain: global sociosexuality (i.e. preferences for casual sex: 79, 80) and specific
measures of mating attitudes and mating behaviors where:

i. Mating attitudes included: preferences for short-term relationships, and sociosexualattitudes and desires.

ii. Mating behaviors included: number of sexual partners, one-night-stands/short-term
relationships, potential conceptions, sociosexual behaviors, extra-pair sex, age at first sexual
intercourse, and number of marital spouses.

- Reproductive domain: including both fertility and reproductive success, described below.

i. Fertility: number of children and grandchildren born, and age at the birth of the firstchild.

ii. Reproductive success: number of surviving children/grandchildren.

- Offspring mortality domain: mortality rate and number of deceased offspring.

Both published and unpublished studies were eligible. We restricted our sample to 252 studies with adult participants (≥ 17 years old). If key variables were collected but the relevant 253 254 analyses were not reported, we contacted authors to request effect sizes or raw data. If data 255 were reported in more than one study, we selected the analysis with the larger sample size or which included appropriate control variables, such as age. Studies using measures that were 256 257 ambiguous and/or not comparable to measures used in other studies were excluded (e.g. measures of body size without information about the proportion of fat/muscle mass, or 258 reproductive data during a very restricted time period). Twin studies where participants were 259 sampled as pairs, population level studies, and studies analyzing both sexes together were also 260 excluded, as well as articles that were not written in English or Swedish as we were not 261 262 sufficiently fluent in other languages to conduct unbiased searching and extraction. Multiple 263 measures from the same study were retained if they met the other criteria.

We chose Pearson's r as our effect size measure and effect sizes not given as r were 264 converted (see Supplementary File 1 for conversion formulas); if effect sizes were not 265 convertible, the study was excluded. Where effect sizes for non-significant results were not 266 stated in the article and could not be obtained, an effect size of 0 was assigned (k = 28). 267 Excluding those effects from the analyses had no effect on any of the results. Twenty-nine 268 percent of all observations (133 of 452, selected randomly) were double coded by the first 269 270 author > 2 months apart. Intracoder agreement was 97%. For coding decisions, see Supplementary Files 2. 271

272

In total, 96 studies were selected (8, 9, 12, 31, 33, 47, 48, 58, 63, 64, 74-76,

81-163), comprising 474 effect sizes from 99 samples and 177,044 unique participants
(Figure 1). This exceeds the number of studies for each of the meta-analyses published
previously (54, 76-78).

276 [INSERT FIGURE 1 AROUND HERE]

277 Statistical analyses

We used the *metafor* package (164) in R 3.6.2 (165). *metafor* transforms Pearson's r to 278 279 Fisher's Z for analysis; for ease of interpretation, effect sizes were converted back to r for presentation of results. For 2D:4D and voice pitch, effects were reverse coded prior to 280 281 analysis because low values denote greater masculinity. Similarly, effects were reverse coded for all offspring mortality outcomes as well as the outcomes age at first birth and age at first 282 sexual intercourse/contact, as low values denote increased fitness. In all analyses reported 283 here, therefore, a positive value of r denotes a positive relationship between masculinity and 284 fitness outcomes. All predicted relationships were positive. 285

Analyses were conducted using random-effects models, as we expected the true effect to vary across samples. We controlled for multiple comparisons by computing q-values (166). Note that q-values estimate the probability that a significant effect is truly significant or not; they are not adjusted p values. Thus, in all analyses presented below, only effects that remained significant after q-value computation (indicated by q-values < .05) are presented as significant. We computed q-values using all p values across all tests conducted in the whole analysis (266 in total). Q-values can be viewed in Supplementary File 7.

The analyses were conducted on three levels for both predictor traits and outcomes (Figure 2). For predictor traits, all six masculine traits were first combined and analyzed together at the *global masculinity level*. At the *trait level*, each masculine trait was then analyzed separately. Lastly, each masculine trait was further divided into separate *trait indices*, which were analyzed as potential moderators (see below).

298 [INSERT FIGURE 2 AROUND HERE]

For the outcomes, mating, reproduction, and offspring mortality were first analyzed 299 together at the *total fitness level*. Given the widespread use of mating measures as proxies of 300 reproductive outcomes, it is imperative where possible to test (and ideally compare) both 301 mating and reproduction, to ensure that we are not relying on proxies that do not measure 302 what they are assumed to measure. The domain level therefore divided outcomes into the 303 mating domain, the reproductive domain, and the offspring mortality domain and analyzed 304 them separately. The last level, the *measures level*, further divided mating and reproduction 305 into their separate measures (mating attitudes and behaviors, and fertility and reproductive 306 success, respectively), which were analyzed as subgroups. 307

The mating domain comprised mating *attitudes* and mating *behaviors*, as high mating success may result from increased mating efforts (reflected in favorable attitudes towards short-term mating) and/or encountering more mating opportunities (reflected in mating behaviors) without actively seeking them (because of female choice, for example). It is therefore necessary to divide these two measures.

The reproductive measures, fertility (number of offspring) and reproductive success (number of surviving offspring), are closely related but were also analyzed separately in subgroup analyses. Offspring mortality, on the other hand, was usually indexed by mortality *rate* (only two studies used absolute number of dead offspring, and it made no difference to the results whether those studies were included or not) and is not directly related to offspring numbers. Offspring mortality was therefore analyzed as a separate domain. As there were too few observations of offspring mortality to test predictor traits separately, this outcome wasonly analyzed at the global masculinity level.

In addition to analyzing all samples together, we also analyzed low and high fertility samples separately to assess whether results were robust in both types of populations. We used a cut-off of three or more children per woman on average within that sample, which roughly corresponds to samples with vs without widespread access to contraception (167). Samples therefore had two levels: *all samples*, and the two sample types *low fertility* and *high fertility*.

The analysis structure was therefore as summarized in Figure 2: *overall analyses* tested global masculinity as a predictor of total fitness, as well as the three domains of mating, reproduction, and offspring mortality, separately, across all samples. In our *main analyses*, we analyzed masculinity at the trait level, in relation to the two outcome domains mating and reproduction. The following *subgroup analyses* considered low and high fertility samples separately, in addition to also dividing outcomes into their respective measures (mating attitudes vs mating behaviors, and fertility vs reproductive success).

Lastly, we performed a series of exploratory meta-regressions on potential moderator 334 335 variables. Such moderation analyses compare effect sizes across categories of studies as determined by a particular study characteristic, e.g. monogamous vs polygynous marriage 336 systems, to determine if effect sizes were robust and/or equivalent across these categories. 337 Since power was often low, we ran moderation analyses separately for each study 338 characteristic rather than trying to test for interactions. For all masculine traits where we had 339 340 sufficient power, trait-general moderation analyses included: domain type (mating vs reproduction), mating measure type (attitudes vs behaviors), reproductive measure type 341 (fertility vs reproductive success), sample type (low vs high fertility), low fertility sample type 342 (student vs non-student), high fertility sample type (traditional vs industrialized), ethnicity, 343

marriage system, publication status (published vs not published effect), peer review status 344 345 (peer reviewed vs not peer reviewed), sexual orientation, transformation of variables, conversion of effect sizes, age control, and inclusion of other control variables. Note that 346 347 since we included many non-published effects from studies that were published, 'publication status' referred to whether particular the particular *effects* were published, not the study as 348 whole. The analysis can therefore detect evidence of any tendency for significant results to be 349 350 'written up' while nonsignificant ones are not, whether this bias occurs between or within manuscripts. We ran moderation analyses both for outcome domains and outcome measures 351 (i.e. mating attitudes and mating behaviors, and fertility and reproductive success, 352 353 respectively). For each masculine trait, we also conducted trait-specific moderation analyses 354 (e.g. subjectively rated vs morphometric facial masculinity (for full details on trait-specific moderators, see Supplementary Files 3). 355

Analyses sometimes included more than one observation from the same study/sample. In all analyses, therefore, effect sizes were clustered both by sample and by study. For all analyses, only relationships with a minimum of three independent samples from a minimum of two separate studies were analyzed. For moderation analyses, this meant that *each category* of the moderator needed observations from at least three samples from at least two studies; in many cases, there were not enough observations to test for moderators.

In the Results section, unless otherwise specified, we summarize results from *traitgeneral* moderation analyses of outcome *domains* only (where results for outcome *measures* and *trait-specific* moderators are reported in Supplementary File 4). Additional details and full results of all analyses can be found in Supplementary Files 3-5.

366

Results

367 Summary of samples

All 96 studies included in the meta-analysis are shown in Figure 1. In total, 29 articles 368 369 reported effect sizes from high fertility samples, which included 17 articles drawing on 13 different extant forager or subsistence populations (of the type sometimes referred to as 'small 370 scale societies', coded here as non-industrialized) predominantly in Africa or Latin America. 371 The remaining high fertility data came from historical samples or low socioeconomic status 372 sub-populations within low-fertility countries (e.g. agricultural Polish communities, former 373 374 'delinquents' in the US, and Zulus living in South African townships). Sixty-nine articles reported data from low fertility populations, which came from 54 primarily student or 375 partially-student samples (43 of which were from English-speaking countries), and only 12 376 377 samples which could be considered representative community or cohort/panel samples. Two articles reported data drawn from 'global' online samples (classified as low fertility). The 378 remaining low fertility samples were either unspecified or sampled particular sub-populations 379 380 (e.g. specific professions).

381 Overall analyses of global masculinity

382 In the initial overall analyses, global masculinity was weakly but significantly associated with greater total fitness (i.e. mating, reproduction, and offspring mortality combined) (r = .080, 383 95% CI: [0.061, 0.101], q = .001; we reiterate here that for all analyses, q-values < .05 denote 384 significance after correcting for multiple comparisons). When we divided the outcome 385 measures into their three domains, the positive (albeit weak) associations with global 386 masculinity remained significant for mating, but not for reproduction or offspring mortality 387 (mating: r = .090, 95% CI: [0.071, 0.110], q = .001; reproduction: r = .047, 95% CI: [0.004, 388 (0.090], q = .080; offspring mortality: r = .002, 95% CI: [-0.011, 0.015], q = .475). While the 389 390 effect was thus only significant for mating, the differences between effects were not significant, but we note that sample sizes differed considerably between domains. 391

Below, we present in further detail the results of the effect of global masculinity on each of the three outcome domains: mating, reproduction, and offspring mortality. We then present the associations between each masculine trait and mating and reproductive measures, separately. We also present results for subgroup and trait-general moderation analyses (for outcome domains only); for complete results, see Supplementary Files 4 and 5.

397 Mating

Main analyses of each masculine trait. This set of analyses tested the prediction that 398 individual masculine traits are positively associated with mating. In terms of the overall 399 mating domain (i.e. mating attitudes and behaviors combined), all masculine traits showed the 400 predicted positive relationships with mating, and the effects were significant for all traits 401 except for facial masculinity and 2D:4D (Table 1). Some of these effects were very weak, 402 403 however. The strongest associations with the mating domain were seen in terms of body masculinity (r = .133, 95% CI: [0.091, 0.176], q = .001; Figure 3), voice pitch (r = .132, 95%404 405 CI: [0.061, 0.204], q = .002; Figure 4), and testosterone levels (r = .093, 95% CI: [0.066, 10.006]406 0.121], q = .001; Figure 5). Height showed a significant but smaller effect size (r = .057, 95%CI: [0.027, 0.087], q = .002; Figure 6). While not the weakest association, the relationship 407 between facial masculinity and mating was nonsignificant (r = .080, 95% CI: [-0.003, 0.164], 408 q = .117). The effect for 2D:4D was also nonsignificant (r = .034, 95% CI: [0.000, 0.069], q =409 .102), and moderation analyses showed that this was the only trait that showed a significantly 410 smaller effect size than the strongest predictor, body masculinity (p < .001, q = .006). 411

412 [INSERT TABLE 1 AROUND HERE]

413

414 *Comparison of high and low fertility samples.* Across all masculine traits, most effect sizes

415 (94%) came from low fertility samples. Moderation analyses of sample type could only be run

416 for body masculinity and height; neither was significant, although in both cases the effect

sizes observed in the main analyses were significant only for low fertility, and not the less 417 418 numerous high fertility samples (k = 4 for each trait). The other four traits had only been measured in one high fertility sample each, and the main analyses thus contained almost 419 420 exclusively low fertility samples. We further compared low fertility samples which were predominantly students with other low fertility samples as part of our moderation analyses 421 where possible, i.e. for body masculinity, voice pitch, height, and testosterone. For body 422 423 masculinity, student samples showed a significantly stronger effect than non-student samples for mating *behaviors* only (B = -.128, p = .009, q = .032) but otherwise we found no 424 differences (see Supplementary File 4). 425

426 [INSERT FIGURES 3, 4, 5 AND 6 AROUND HERE]

427

Inclusion bias/heterogeneity. Since the analysis included unpublished data, the distribution of effects in the funnel plots (see Supplementary File 6A) shows availability bias rather than publication bias. Apart from voice pitch, for which we did not have many effects, visual inspection of funnel plots indicated that they were generally symmetric, suggesting that the analysis did not systematically lack studies with unexpected small effects. There was significant heterogeneity of effect sizes for facial masculinity, body masculinity, and height; all of which are accounted for in a random-effects analysis.

Additional subgroup and moderation analyses for outcome domains. In this step of the
analyses, we tested the hypothesis that each of the six masculine traits is positively associated
with the two mating domain measures (mating attitudes and mating behaviors) and tested
further potential control variables and trait-specific moderators. Results of subgroup analyses
can be viewed in Table 1 and trait-general moderators in Table 3; full results of all
moderation analyses are reported in Supplementary Files 4.

Type of mating measure (attitudes vs behaviors) was never a significant moderator. 441 442 However, for both body masculinity and height, there were significant effects for mating behaviors (body masculinity: r = .142, 95% CI: [0.099, 0.187], q = .001, height: r = .054, 95%443 CI: [0.021, 0.087], q = .008) but not attitudes. Voice pitch was significantly related to mating 444 behaviors (r = .124, 95% CI: [0.043, 0.206], q = .016) but was not measured in combination 445 with mating attitudes. Testosterone levels showed near identical effects for both mating 446 attitudes and behaviors (r = .099, 95% CI: [0.026, 0.173], q = .032 and r = .084, 95% CI: 447 [0.058, 0.110], q = .001, respectively). 448

No trait-general moderator consistently changed the pattern of the associations (Table 449 450 3). Body masculinity effects were stronger in studies where age had not been controlled for compared to where it had been controlled for (B = 0.103, p = .015, q = .047). Associations for 451 2D:4D were weaker in non-white/mixed ethnicity samples compared to white samples (B = -452 0.080, p = .014, q = .047), and stronger where variables had been transformed to approximate 453 normality compared to when they had not been transformed (B = 0.103, p = .016, q = .047). 454 455 Similarly, associations for testosterone levels were also stronger for normality-transformed variables (B = 0.057, p = .015, q = .047), and weaker in gay/mixed sexuality samples 456 compared to in heterosexual samples (B = -0.059, p = .003, q = .016). 457

For trait-specific moderators, significant moderation was seen for *type* of body masculinity where body shape was a significantly weaker predictor than strength (B = -0.099, p = .003, q = .017). Effects for *rated* body masculinity were significantly stronger than for indices taken from body measurements (B = 0.177, p = .007, q = .029). For 2D:4D, studies that had measured digit ratios three times – rather than twice or an unknown number of times – showed significantly stronger effects (B = 0.102, p = .006, q = .025).

464 **Reproduction**

465 *Main analyses of each masculine trait.* In this set of analyses, we tested the hypothesis that 466 individual masculine traits positively predict reproduction. As Table 2 shows, relationships 467 were generally in the predicted direction, but body masculinity was the strongest and only 468 significant predictor (r = .143, 95% CI: [0.076, 0.212], q = .001; Figure 7). The only trait with 469 an effect size significantly smaller than body masculinity was height (B = -0.107, p = .005, q470 = .023).

471 [INSERT TABLE 2 AROUND HERE]

472 [INSERT FIGURE 7 AROUND HERE – FIGURE CAPTION BELOW]

473

Comparison of high and low fertility samples. The majority (77 %) of observations of 474 reproduction were from high fertility samples. Moderation analyses of low versus high 475 fertility samples could only be conducted for 2D:4D and height; effect sizes did not differ 476 significantly between sample types. Comparing types of high fertility samples (industrialized 477 vs non-industrialized) for 2D:4D and height did not show any differences in effect sizes (see 478 479 Supplementary File 4). It was not possible to compare sample subtypes for the other traits 480 because observations were almost entirely from non-industrialized populations. 481 Inclusion bias/heterogeneity. Visual inspection of funnel plots (see Supplementary File 6B) 482 suggested that while the effects for voice pitch, height, and testosterone levels were 483 symmetrically distributed, our analysis may have lacked studies for the other traits. Facial 484 masculinity and height showed significant heterogeneity. 485 486

487 [INSERT TABLE 3 AROUND HERE]

488

Additional subgroup and moderator analyses for outcome domains. Results of subgroup
analyses can be viewed in Table 2 and trait-general moderators in Table 3; full results of
moderation analyses are found in Supplementary File 4.

Moderation analyses (where possible) showed no evidence that the effects of masculinity traits on fertility differed from the effects on reproductive success. However, for body masculinity, the effect on fertility was significant (r = .130, 95% CI: [0.060, 0.201], q =.002; five out of six samples high fertility) while the somewhat larger effect on reproductive success was not. For 2D:4D, there was a significant effect for reproductive success (four out of five samples from high fertility populations: r = .174, 95% CI: [0.085, 0.267], q = .002) but not for fertility.

Similarly, for mating, no trait-general or trait-specific moderators had any consistent effects on the results. Body masculinity effects were stronger where effect sizes had been converted to Pearson's *r* compared to where they initially had been given as r (B = 0.143, p =.015, q = .047), and effects for height were stronger in gay/mixed sexuality samples than heterosexual samples (B = 0.135, p = .016, q = .047).

504 Comparing mating and reproduction across traits

Moderation analyses of domain type (mating versus reproduction) for each trait showed no significant differences, although height and testosterone levels had weaker associations with reproduction than mating while body masculinity showed the opposite pattern. There were generally far fewer observations for reproductive measures, so this nonsignificant analysis may reflect lack of power. For facial masculinity, voice pitch, and 2D:4D, effect sizes for global mating and reproductive measures were near identical.

512

Discussion

513 Summary of results

We conducted the first comprehensive meta-analysis of the relationships between men's 514 515 masculine traits and outcomes related to mating and reproduction. Various proposed (and non-mutually exclusive) hypotheses suggest that more masculine men should show increased 516 517 mating success (indexed by more matings and/or preferences for short-term mating), increased reproductive output (indexed by fertility and/or reproductive success), and/or lower 518 offspring mortality. Our results showed partial support for these predictions. Global 519 520 masculinity (i.e. all masculine traits combined) significantly predicted effects in the mating domain, but not the reproductive domain or the offspring mortality domain. When we 521 analyzed each masculine trait separately, all traits except facial masculinity and 2D:4D 522 523 significantly predicted effects in the mating domain, where similarly strong associations were 524 seen for body masculinity, voice pitch, and testosterone levels, and a weaker correlation was seen for height. In terms of the reproductive domain, the only significant predictor was body 525 526 masculinity. It was not possible to analyze offspring mortality at the specific predictor level owing to a severe lack of relevant data from which to draw conclusions (total number of 527 observations for each outcome domain: mating domain k = 371; reproductive domain k = 81; 528 offspring mortality domain k = 22). 529

We also examined how these effects play out in high versus low fertility populations. 530 Typically, however, different outcomes were measured in different groups of populations; 531 mating outcomes were predominantly measured in low fertility populations, while 532 reproductive outcomes were measured mainly in high fertility populations. This made it more 533 challenging to draw direct comparisons. Where it was possible to run moderation analyses on 534 535 sample type, there were no significant differences. These analyses, however, have small 536 numbers of high and low fertility samples in mating and reproductive outcomes respectively. Therefore, while we can confidently say that most forms of masculinity (but not facial 537 masculinity or 2D:4D) are associated with (largely self-reported) mating outcomes in low 538

fertility samples, we cannot draw any clear conclusions regarding mating success in high
fertility samples. Similarly, although we are confident that body masculinity is associated
with fertility/reproductive success in high fertility samples, we cannot draw conclusions about
low fertility contexts.

More generally, our moderation analyses on outcome types and factors relating to 543 measure quality did not yield any consistent differences between effect sizes, suggesting that 544 545 the effects we do find are reasonably robust within sample type at least. Two key points to note here are that: *i*. although effect sizes for mating attitudes and mating behaviors did differ 546 for some traits (i.e. facial masculinity and body masculinity), these differences were never 547 548 significant, despite mating behaviors being constrained by opportunities (assuming participants report truthfully), and *ii*. similarly, effect sizes did sometimes differ by 549 publication status but never significantly so; in addition, the direction of the differences was 550 551 not consistent (i.e. effect sizes were not consistently larger in published analyses). Even if the analysis was restricted to nonpublished effects only, the association between body masculinity 552 553 and both mating and reproduction would be weaker but remain significant (mating: r = .077, p = .006; reproduction: r = .112, p < .001; both associations would remain significant after q-554 value computation). Overall, this suggests that researchers have not been selectively reporting 555 larger effect sizes. 556

557 Compared to previous meta-analyses assessing associations between handgrip strength 558 and mating outcomes (76), height/strength and reproductive outcomes (77, 78), and 559 testosterone levels and mating effort (54), our analysis benefits from more comprehensive 560 measures of masculinity, larger sample sizes, and inclusion of more unpublished effects. With 561 the exception of Xu and colleagues' analysis (78), we observe smaller effect sizes than 562 previous meta-analyses, which suggests that the association between masculinity and fitness 563 outcomes has previously been overestimated. In general, what significant associations we did observe were small and ranged between r = .05 and .17, although they are potentially meaningful in an evolutionary context. As benchmarks for interpreting correlations, Funder and Ozer (168) suggest that a correlation of .10, while being a small effect, has the potential to be influential over a long time period, and a medium-size correlation of .20 can be consequential both in the short- and long-term. The cumulative effect of relatively 'weak' correlations can therefore be of real consequence, particularly when considered in terms of selection acting over many generations.

571

572 Major implications

Selection for body masculinity. The first stand-out result of our analysis is that body
masculinity (i.e. strength/muscularity) is the only trait in our analysis that was consistently
correlated with both mating and reproductive outcomes across populations, and the effects of
body masculinity on these outcomes were among the strongest in the analysis. In contrast,
other aspects of masculinity (except facial masculinity and 2D:4D) predicted mating success
in low fertility samples but did not yield reproductive benefits in high fertility samples.

Body masculinity is therefore the trait where we have the most compelling evidence 579 that selection is currently happening within naturally fertile populations - and from that, can 580 581 infer that selection likely took place in prior eras as well. As such, our results are consistent with the argument that dimorphisms in strength and muscle mass are sexually selected. 582 Overall, since traits such as body size, strength, and muscularity are associated with 583 584 formidability, our findings are consistent with the male-male competition hypothesis. In species with male intrasexual competition, males tend to evolve to become larger, stronger, 585 and more formidable than females, as they are in humans. Some authors argue that male-male 586 violence has influenced human evolution (35, 169), and male intergroup aggression increases 587 mating/reproductive success in both non-industrialized human societies and in non-human 588

primates (170, 171). (And indeed the non-human evidence might suggest this form of 589 590 dimorphism has been under selection since pre-hominid ancestors, although the strength of such selection pressures have likely fluctuated over this time [172].) For example, in the 591 Yanomamö Indians, men who kill others have greater reproductive success (173). A 592 relationship between formidable traits and fitness outcomes need not be a direct one, however. 593 It might, as mentioned in the introduction, be mediated by other factors that are important in 594 595 mate choice, such as interpersonal status and dominance. For example, features that are advantageous in intraspecies conflicts may also be advantageous when hunting game (37); 596 Smith and colleagues (144) reported that in a hunter-gatherer population, men with greater 597 598 upper body strength and a low voice pitch had increased reproductive success, but this relationship was explained by hunting reputation. 599

It is of course possible that different selection pressures may have contributed to the evolution of different masculine traits. Male-male competition for resources and mates, female choice, and intergroup violence are all plausible, non-mutually exclusive explanations (172). In this article, we have focused on the effect of men's own traits on their fitness, but it is of course equally possible that men varying in masculinity may differ in the quality of the mates they acquire. If masculine men are able to secure mates who are more fertile and/or better parents, this may also increase their fitness.

607

No evidence of advantage for facial masculinity. Considerable attention has been given in the literature to the hypothesis that masculinity in men's facial structure is an indicator of heritable immunocompetence (i.e. good genes), which should then be associated with greater mating and reproductive success. While we find that the effect of facial masculinity on mating was similar in size to that of other traits (r = .08), it was not significantly different from zero, suggesting more variability in effects. Furthermore, the effect of facial masculinity on mating

(such as it was) was largely driven by mating attitudes and was close to zero for mating 614 615 behaviors, suggesting that men's facial masculinity exerts virtually no influence on mating when moderated by female choice. Similarly, the influence of facial masculinity on fertility in 616 617 high fertility samples was non-existent (r = .00). Although the relationship with reproductive success appeared stronger, this was based on only two samples. This is, all together, doubly 618 striking because although voice pitch, height, and testosterone levels did not predict 619 620 reproductive outcomes, they did all relate to mating in the expected direction. Facial masculinity is ergo an outlier in being so entirely unrelated to mating success in our data, 621 while subject to so large a literature assuming the opposite. 622

Overall, these findings contradict a large body of literature claiming that women's preferences for masculinity in men's faces are adaptive. Rather, they indicate that such preferences (to the extent they exist at all) are a modern anomaly only found in industrialized populations, as suggested by Scott and colleagues (45), and as demonstrated by the positive correlation between facial masculinity preferences and national health and human development indices (27).

629

Students and foragers. One key observation regarding our dataset is that it shows a rather 630 'bimodal' distribution between a large number of studies sampling (predominantly English-631 speaking) students on one hand, and a cluster of studies sampling foragers, horticulturalists, 632 and other subsistence farmers (predominantly from just two continents) on the other. Where it 633 634 was possible to compare student vs non-student/mixed samples within low fertility populations, and traditional vs industrialized high fertility samples, we generally did not find 635 636 any differences. Likewise, where it was possible to compare monogamous and formally polygynous cultures, we also found no differences. This is despite evidence that monogamy 637 actually changes selection pressures on human men (61). Therefore, although we are 638

reasonably confident that our results regarding body masculinity and reproduction are robust, insofar as they are based on non-industrialized populations with a range of subsistence patterns (hunter-gatherers, forager-horticulturalists, and pastoralists), it remains essential to consider rebalancing the literature. Not only do we require more holistic representation of non-industrialized populations (drawing from Asia and Oceania in particular, where we had one and zero samples, respectively), but it is also important to increase representation of nonstudent participants in low fertility contexts.

646

Disconnection between mating and reproductive literatures. As noted above, we found that 647 648 voice pitch, height, and testosterone levels were associated with (largely self-reported) mating success in mostly low fertility populations, but not with actual reproductive fitness in high 649 fertility populations. A caveat here is that effect sizes for voice pitch and reproduction were 650 651 similar in strength to effect sizes for body masculinity, but we note that this analysis had the smallest sample size of our whole analysis (k = 5, n = 143), which prevents us from drawing 652 firm conclusions regarding the relationship between voice pitch and reproductive outcomes. 653 Overall, however, the contradicting pattern of results for the traits mentioned above 654 raise important concerns for the human sexual selection field, particularly with respect to 655 656 whether (and which) mating measures can be used as reliable indicators of likely ancestral fitness when considering the current evidence base. Since reproductive outcomes - for good 657 reason - are not considered meaningful fitness measures in populations with widespread 658 contraception use, we typically test fitness outcomes in industrialized populations using 659 mating measures such as sociosexual attitudes and casual sexual encounters. This is done 660 661 under the assumption that such measures index mating strategies that ancestrally would have increased men's offspring numbers. However, if mating outcomes (be it attitudinal or 662 behavioral) measured in low fertility populations truly index reproductive outcomes in 663

naturally fertile contexts, we would expect traits that predict mating to also predict
reproduction on average across samples (notwithstanding the diversity in norms/reproductive
behaviors across high fertility samples). We do not, however, have evidence that this is
generally the case. Our findings therefore raise the question of whether these widely used
measurements are truly valid proxies of what we purport to be measuring.

669 Our findings thus illustrate that when we attempt to test the same underlying research 670 questions using different measurements in different populations, this may yield conclusions 671 that are erroneous or misleading when applied outside of the studied population. We suggest, 672 based on our analysis, that researchers could for instance consistently gather sexual partner 673 number, age of marriage, and number/survival rates of offspring in multiple population types. 674 Wherever possible, it is essential to use the same measurements across populations, or at least 675 resist the temptation of applying our findings universally.

676

677 Key limitations

Non-linearity. A limitation of our analysis is that we only assessed linear relationships, 678 ignoring possible curvilinear associations. There is evidence suggesting that moderate levels 679 of masculinity might be associated with increased reproductive success (see e.g. 63, for 680 681 offspring survival rates) and perceived attractiveness (31, 174, but see also 34), with a decrease for both very low and very high levels of masculinity. Indeed, some of these authors 682 have argued that masculinity may be under stabilizing, rather than directional, selection in 683 humans. In instances such as these, our 'null' conclusions regarding e.g. facial masculinity, 684 remain valid; facial masculinity does not appear to be under directional selection. However, 685 we also note that there is data suggesting that height in men may be optimal when it is over-686 average but not maximal. In this scenario, although the linear relationship would be weaker, 687 the trait remains under directional selection, and we would still expect to see positive, albeit 688

weak, associations in our analyses. In the vast majority of studies included, only linear
relationships were tested, and acquiring original data to investigate and synthesize non-linear
effects was beyond the scope of the current article. However, increased publication of open
data with articles may well facilitate such a project in future years.

693

Testosterone effects. As mentioned above, in our analysis testosterone levels predicted mating 694 695 outcomes - with similar effect sizes for attitudinal and behavioral measures - but did not predict reproduction. While a causal relationship between testosterone levels and mating 696 success cannot be established from this (i.e. whether high testosterone men pursue more 697 698 mating opportunities which leads to more matings, or whether high testosterone results from 699 many matings), testosterone is commonly argued to motivate investment in mating effort. If 700 current testosterone levels index degree of masculine trait expression in men, our results 701 *might* indicate that masculine men's increased mating success is due to greater pursuit of matings - rather than reflecting female choice and/or greater competitiveness. Two caveats for 702 703 interpreting our results, however (applicable both to the significant effect we observe for mating and the nonsignificant effect for reproduction), is that circulating testosterone levels *i*. 704 change over the course of a man's lifetime, peaking in early adulthood and subsequently 705 706 declining (50; although this may not be the case in non-industrialized populations: 175), and *ii.* are reactive. In the studies we gathered, testosterone levels were generally measured 707 contemporaneously with mating/reproductive data collection - not when masculine traits 708 709 generally become exaggerated in adolescence. Testosterone also decreases, for example, when men enter a relationship or get married (176, 177), when they become fathers (176, 178), or 710 when they engage in childcare (176). Thus, men whose testosterone levels were previously 711 high may show declining testosterone levels either because of their age and/or because their 712 relationship or fatherhood status has changed. This limits the conclusions we can draw, both 713

with regards to a potential mediating role of testosterone levels in the association between masculine traits and mating success, and the observed nonexistent effect for testosterone levels and reproductive outcomes. We also note that the sample size for reproduction, as a function of testosterone levels, was small.

718

719 Conclusion

720 In summary, we used a large-scale meta-analysis of six masculine traits and their relationships with mating and reproductive outcomes to test whether such traits are currently 721 under selection in humans. We found that all masculine traits except facial masculinity and 722 723 2D:4D were associated with significantly greater mating success. However, only body masculinity predicted higher fertility, indexed by reproductive onset, number of offspring, and 724 grand-offspring. We further note that the mating and reproduction literature is starkly split 725 726 between studying mating in predominantly student settings, and 'only' fertility in high fertility settings, which imposes constraints on both this paper and our field as a whole. We argue that 727 728 our findings illustrate that when we test hypotheses about human evolution largely in industrialized populations, we risk drawing conclusions that are not supported outside of 729 730 evolutionarily novel, highly niche mating and reproductive contexts. We therefore call for 731 greater sample diversity and more homogenous measurements in future research.

732

733 **References**

J. M. Plavcan, Sexual dimorphism in primate evolution. *Am. J. Phys. Anthropol.* 116(S33), 25-53 (2001).

J. H. Crook, "Sexual selection, dimorphism, and social organization in the primates" in
 Sexual selection and the descent of man, B. G. Campbell, Ed. (Aldine Publishing
 Company, 1972).

739	3.	M. F. Hammer, F. L. Mendez, M. P. Cox, A. E. Woerner, J. D. Wall, Sex-biased
740		evolutionary forces shape genomic patterns of human diversity. PLoS Genet 4(9),
741		e1000202 (2008).

- 4. J. C. Mitani, J. Gros-Louis, A. F. Richards, Sexual dimorphism, the operational sex ratio,
 and the intensity of male competition in polygynous primates. *Am. Nat.* 147(6), 966-980
 (1996).
- J. P. Swaddle, G. W. Reierson, Testosterone increases perceived dominance but not
 attractiveness in human males. *Proc. Royal Soc* 269(1507), 2285-2289 (2002).
- 747 6. J. P. Gray, L. D. Wolfe, Height and sexual dimorphism of stature among human societies.
 748 *Am. J. Phys. Anthropol.* 53(3), 441-456 (1980).
- 749 7. R. J. Smith, W. L. Jungers, Body mass in comparative primatology. *J. Hum. Evol.* 32(6),
 523-559 (1997).
- 8. W. D. Lassek, S. J. C. Gaulin, Costs and benefits of fat-free muscle mass in men:
- relationship to mating success, dietary requirements, and native immunity. *Evol. Hum. Behav.* 30(5), 322-328 (2009).
- 9. S. M. Hughes, G. G. Gallup, Sex differences in morphological predictors of sexual
- behavior: Shoulder to hip and waist to hip ratios. *Evol. Hum. Behav.* 24(3), 173-178(2003).
- 10. D. Singh, Adaptive significance of female physical attractiveness: Role of waist-to-hip
 ratio. *J. Pers. Soc. Psychol.* 65(2), 293-307 (1993).
- 11. M. Tovée, D. Maisey, E. Vale, P. Cornelissen, Characteristics of male attractiveness for
 women. *Lancet* 353(9163), 1500-1500 (1999).
- 12. J. Weeden, J. Sabini, Subjective and objective measures of attractiveness and their relation
 to sexual behavior and sexual attitudes in university students. *Arch. Sex. Behav.* 36(1), 7988 (2007).

- 764 13. J. T. Manning, *Digit ratio: A pointer to fertility, behavior, and health.* (Rutgers University
 765 Press, 2002).
- 14. C. L. Apicella, V. A. Tobolsky, F. W. Marlowe, K. W. Miller, Hadza hunter-gatherer men
 do not have more masculine digit ratios (2D:4D). *Am. J. Phys. Anthropol.* 159(2), 223-232
 (2016).
- 15. D. A. Puts, C. L. Apicella, R. A. Cárdenas, Masculine voices signal men's threat potential
 in forager and industrial societies. Proc. Royal Soc. B. 279(1728), 601-609 (2012).
- 16. F. Galis, C. Ten Broek, S. Dongen, L. Wijnaendts, Sexual dimorphism in the prenatal
- digit ratio (2D:4D). Arch. Sex. Behav. 39(1), 57-62 (2010).
- 17. G. Richards, M. Gomes, T. Ventura, Testosterone measured from amniotic fluid and
- maternal plasma shows no significant association with directional asymmetry in newborn

digit ratio (2D:4D). J. Dev. Orig. Health Dis. 10(3), 362-367 (2019).

- 18. S. A. Butterfield, R. A. Lehnhard, E. M. Loovis, T. Coladarci, D. Saucier, Grip strength
- performances by 5- to 19-year-olds. *Percept Mot Skills* 109(2), 362-370 (2009).
- P. Y. Fechner, "The biology of puberty: New developments in sex differences" in *Gender differences at puberty*, C. Hayward, Ed. (Cambridge University Press, 2003).
- 20. E. M. Weston, A. E. Friday, P. Liò, Biometric evidence that sexual selection has shaped
 the hominin face. *PLoS One* 2(8), e710. (2007).
- 782 21. I. Folstad, A. J. Karter, Parasites, bright males, and the immunocompetence handicap. *Am.*783 *Nat.* 139(3), 603-622 (1992).
- 22. M. P. Muehlenbein, R. G. Bribiescas, Testosterone-mediated immune functions and male
 life histories. *Am. J. Hum. Biol.* 17(5), 527-558 (2005).
- 786 23. L. G. Boothroyd, B. C. Jones, D. M. Burt, R. E. Cornwell, A. C. Little, B. P. Tiddeman,
- 787 D. I. Perrett, Facial masculinity is related to perceived age, but not perceived health. *Evol*.
- 788 *Hum. Behav.* 26(5) 417-431 (2005).

789 24. J. Nowak, B. Pawłowski, B. Borkowska, D. Augustyniak, Z. Drulis-Kawa, No evidence
790 for the immunocompetence handicap hypothesis in male humans. *Sci. Rep.* 8(1), 1-11

791 (2018).

- 25. L. G. Boothroyd, I. Scott, A. W. Gray, C. I. Coombes, N. Pound, Male facial masculinity
 as a cue to health outcomes. *Evol. Psychol.* 11(5) 1044-58 (2013).
- 794 26. Y. Z. Foo, L. W. Simmons, D. I. Perrett, P. G. Holt, P. R. Eastwood, G. Rhodes, Immune
- function during early adolescence positively predicts adult facial sexual dimorphism in
 both men and women. *Evol. Hum. Behav.* 41(3), 199-209 (2020).
- 797 27. U. M. Marcinkowska, M. J. Rantala, A. J. Lee, M. V. Kozlov, T. Aavik, H. Cai, ... B. J.
- Dixson, Women's preferences for men's facial masculinity are strongest under favorable
 ecological conditions. *Sci. Rep.* 9(1), 1-10 (2019).
- 28. I. M. L. Scott, A. P. Clark, L. G. Boothroyd, I. S. Penton-Voak, Do men's faces really
 signal heritable immunocompetence? *Behav. Ecol.* 24(3), 579-589 (2013).
- 29. A. A. Zaidi, J. D. White, B. C. Mattern, C. R. Liebowitz, D. A. Puts, P. Claes, M. D.
- 803 Shriver, Facial masculinity does not appear to be a condition-dependent male ornament
- and does not reflect MHC heterozygosity in humans. *Proc. Natl. Acad. Sci. U.S.A.* 116(5),
 1633 (2019).
- 30. A. C. Little, "Attraction and human mating" in *Evolutionary perspectives on social*
- 807 *psychology*, V. Zeigler-Hill, L. L. Welling, T. K. Shackelford, Eds. Springer, Cham.
- 808 (2015) pp. 319-332.
- 31. D. A. Frederick, M. G. Haselton, Why is muscularity sexy? Tests of the fitness indicator
 hypothesis. *Pers. Soc. Psychol. Bull.* 33(8), 1167-1183 (2007).
- 811 32. P. B. Gray, D. A. Frederick, Body image and body type preferences in St. Kitts,
- 812 Caribbean: A cross-cultural comparison with US samples regarding attitudes towards
- 813 muscularity, body fat, and breast size. *Evol. Psychol.* 10(3), 147470491201000319 (2012).

- 33. A. W. Lukaszewski, C. M. Larson, K. A. Gildersleeve, J. R. Roney, M. G. Haselton,
- 815 Condition-dependent calibration of men's uncommitted mating orientation: Evidence
 816 from multiple samples. *Evol. Hum. Behav.* 35(4), 319-326 (2014).
- 817 34. A. Sell, A. W. Lukazsweski, M. Townsley, Cues of upper body strength account for most
- of the variance in men's bodily attractiveness. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*
- 819 284(1869), 20171819 (2017).
- 35. A. K. Hill, D. H. Bailey, D. A. Puts, "Gorillas in our midst? Human sexual dimorphism
 and contest competition in men" in *On human nature*, E. O. Wilson, (Academic Press,
 2016).
- 823 36. D. Puts, Human sexual selection. *Curr Opin Psychol* 7, 28-32 (2016).
- 37. A. Sell, L. Hone, N. Pound, The importance of physical strength to human males. *Hum. Nat* 23(1), 30-44 (2012).
- 38. A. Sell, L. Cosmides, J. Tooby, D. Sznycer, C. von Rueden, M. Gurven, Human
- adaptations for the visual assessment of strength and fighting ability from the body and

828 face. *Proc. Biol. Sci.* 276(1656), 575-584 (2009).

- 39. P. K. Durkee, A. T. Goetz, A. W. Lukaszewski, Formidability assessment mechanisms:
- Examining their speed and automaticity. *Evol. Hum. Behav.* 39(2), 170-178 (2018).
- 40. M. L. Butovskaya, S. Windhager, D. Karelin, A. Mezentseva, K. Schaefer, B. Fink,
- Associations of physical strength with facial shape in an African pastoralist society, the
- 833 Maasai of Northern Tanzania. *PLoS One* 13(5), e0197738 (2018).
- 41. M. P. Haselhuhn, M. E. Ormiston, E. M. Wong, Men's facial width-to-height ratio
 predicts aggression: A meta-analysis. *PLoS One* 10(4), e0122637 (2015).
- 42. J. Raine, K. Pisanski, A. Oleszkiewicz, J. Simner, D. Reby, Human listeners can
- accurately judge strength and height relative to self from aggressive roars and speech.
- *iScience* 4, 273-280 (2018).

- 43. A. C. Little, V. Třebický, J. Havlíček, S. C. Roberts, K. Kleisner, Human perception of
 fighting ability: Facial cues predict winners and losers in mixed martial arts fights. *Behav. Ecol.* 26(6), 1470-1475 (2015).
- 44. D. A. Puts, T. Aung, Does men's voice pitch signal formidability? A reply to Feinberg et
 al. *Trends Ecol. Evol.* 34(3), 189-190 (2019).
- 45. I. M. Scott, A. P. Clark, S. C. Josephson, A. H. Boyette, I. C. Cuthill, R. L. Fried, M. A.
- Gibson, B. S. Hewlett, M. Jamieson, W. Jankowiak, P. L. Honey, Z. Huang, M. A.
- Liebert, B. G. Purzycki, J. H. Shaver, J. J. Snodgrass, R. Sosis, L. S. Sugiyama, V.
- 847 Swami, D. W. Yu, Y. Zhao, I. S. Penton-Voak, Human preferences for sexually dimorphic
- faces may be evolutionarily novel. *Proc. Natl. Acad. Sci. U.S.A.* 111(40), 14388-14393
- 849 (2014).
- 46. P. J. Weatherhead, R. J. Robertson, Offspring quality and the polygyny threshold: "The
 sexy son hypothesis". *Am. Nat.* 113(2), 201-208 (1979).
- 47. A. K. Hill, J. Hunt, L. L. M. Welling, R. A. Cardenas, M. A. Rotella, J. R. Wheatley, K.
- Dawood, M. D. Shriver, D. A. Puts, Quantifying the strength and form of sexual selection
 on men's traits. *Evol. Hum. Behav.* 34(5), 334-341 (2013).
- 48. T. L. Kordsmeyer, J. Hunt, D. A. Puts, J. Ostner, L. Penke, The relative importance of
 intra- and intersexual selection on human male sexually dimorphic traits. *Evol. Hum.*
- 857 *Behav.* 39(4), 424-436 (2018).
- 49. R. B. Slatcher, P. H. Mehta, R. A. Josephs, Testosterone and self-reported dominance
- 859 interact to influence human mating behavior. Soc. Psychol. Pers. Sci. 2(5), 531-539
- 860 (2011).
- 50. A. Booth, J. M. Dabbs Jr, Testosterone and men's marriages. *Soc Forces*. 72(2), 463-477
 (1993).

- 51. L. G. Boothroyd, B. C. Jones, D. M. Burt, D. I. Perrett, Partner characteristics associated
 with masculinity, health and maturity in male faces. *Pers. Individ. Differ.* 43(5), 11611173 (2007).
- 52. M. N. Muller, F. W. Marlowe, R. Bugumba, P. T. Ellison, Testosterone and paternal care
- in East African foragers and pastoralists. *Proc. Royal Soc. B.* 276(1655), 347-354 (2009).
- 868 53. C. Schild, J. Stern, L. Penke, I. Zettler, Voice pitch A valid indicator of one's
- unfaithfulness in committed relationships?. *Adapt. Hum. Behav. Physiol.* 1-16 (2020).
- 54. N. M. Grebe, R. E. Sarafin, C. R. Strenth, S. Zilioli, Pair-bonding, fatherhood, and the role
- of testosterone: A meta-analytic review. *Neurosci Biobehav Rev* 98, 221-233 (2019).
- 55. C. T. Halpern, J. R. Udry, B. Campbell, C. Suchindran, Testosterone and pubertal
- development as predictors of sexual activity: A panel analysis of adolescent males.
- 874 *Psychosom. Med.* 55(5), 436 (1993).
- 56. I. van Bokhoven, S. H. M. van Goozen, H. van Engeland, B. Schaal, L. Arseneault, J. R.
- 876 Séguin, J. M. Assaad, D. S. Nagin, F. Vitaro, R. E. Tremblay, Salivary testosterone and
- aggression, delinquency, and social dominance in a population-based longitudinal study of
- adolescent males. *Horm. Behav.* 50(1), 118-125 (2006).
- 57. C. E. Lefevre, G. J. Lewis, D. I. Perrett, L. Penke, Telling facial metrics: Facial width is
 associated with testosterone levels in men. *Evol. Hum. Behav.* 34(4), 273-279 (2013).
- 58. M. Peters, L. W. Simmons, G. Rhodes, Testosterone is associated with mating success but
 not attractiveness or masculinity in human males. *Anim. Behav.* 76, 297-303 (2008).
- 59. S. W. Gangestad, J. A Simpson, The evolution of human mating: Trade-offs and strategic
 pluralism. *Behav. Brain Sci.* 23(4), 573-587 (2000).
- 60. D. Pérusse, Cultural and reproductive success in industrial societies: Testing the
- relationship at the proximate and ultimate levels. *Behav Brain Sci* 16(2), 267-283 (1993).

- 61. G. R. Brown, K. N. Laland, M. B. Mulder, Bateman's principles and human sex roles. *Trends Ecol. Evol.* 24(6), 297-304 (2009).
- 62. M. Borgerhoff Mulder, Serial monogamy as polygyny or polyandry?: Marriage in the
 Tanzanian Pimbwe. *Hum. Nat.* 20(2), (2009).
- 63. L. G. Boothroyd, A. W. Gray, T. N. Headland, R. T. Uehara, D. Waynforth, D. M. Burt,
- 892 N. Pound, Male facial appearance and offspring mortality in two traditional societies.

893 *PLoS One* 12(1), e0169181 (2017).

64. B. Pawlowski, L. G. Boothroyd, I. D. Perrett, S. Kluska, Is female attractiveness related to

final reproductive success?. *Coll. Antropol.* 32(2), 457-460 (2008).

- 65. M. D. Jennions, A. T. Kahn, C. D. Kelly, H. Kokko, Meta-analysis and sexual selection:
- Past studies and future possibilities. *Evol. Ecol.* 26(5), 1119-1151 (2012).
- 66. S. Van Dongen, S. W. Gangestad, Human fluctuating asymmetry in relation to health and
 quality: A meta-analysis. *Evol. Hum. Behav.* 32(6), 380-398 (2011).
- 900 67. K. Gildersleeve, M. G. Haselton, M. R. Fales, Do women's mate preferences change
- across the ovulatory cycle? A meta-analytic review. *Psychol. Bull.* 140(5), 1205 (2014).
- 68. W. Wood, L. Kressel, P. D. Joshi, B. Louie, Meta-analysis of menstrual cycle effects on
 women's mate preferences. *Emot Rev.* 6(3), 229-249 (2014).
- 904 69. B. C. Jones, A. C. Hahn, C. I. Fisher, H. Wang, M. Kandrik, C. Han, ... L. M. DeBruine,
 905 No compelling evidence that preferences for facial masculinity track changes in women's

906 hormonal status. *Psychol. Sci.* 29(6), 996-1005 (2018).

- 907 70. J. Jünger, N. V. Motta-Mena, R. Cardenas, D. Bailey, K. A. Rosenfield, C. Schild, ... D.
- A. Puts, Do women's preferences for masculine voices shift across the ovulatory cycle?.
- 909 *Horm Behav.* 106, 122-134 (2018).

- 910 71. U. M. Marcinkowska, G. Jasienska, P. Prokop, A comparison of masculinity facial
- 911 preference among naturally cycling, pregnant, lactating, and post-menopausal women.

912 Arch. Sex. Behav. 47(5), 1367-1374 (2018).

- 913 72. C. R. Harris, H. Pashler, L. Mickes, Elastic analysis procedures: An incurable (but
- 914 preventable) problem in the fertility effect literature. Comment on Gildersleeve, Haselton,
- 915 and Fales (2014). *Psychol. Bull.* 140(5), 1260–1264 (2014).
- 73. K. Gildersleeve, M. G. Haselton, M. R. Fales, Meta-analyses and p-curves support robust
 cycle shifts in women's mate preferences: Reply to Wood and Carden (2014) and Harris,

918 Pashler, and Mickes (2014). *Psychol. Bull.* 140(5), 1272–1280 (2014).

- 919 74. S. Arnocky, J. M. Carré, B. M. Bird, B. J. P. Moreau, T. Vaillancourt, T. Ortiz, N. Marley,
- 920 The facial width-to-height ratio predicts sex drive, sociosexuality, and intended infidelity.
- 921 Arch. Sex. Behav. 47(5), 1375-1385 (2018).
- 922 75. G. Rhodes, L. W. Simmons, M. Peters, Attractiveness and sexual behavior: Does
- attractiveness enhance mating success? *Evol. Hum. Behav.* 26(2), 186-201 (2005).
- 924 76. S. Van Dongen, E. Sprengers, Hand grip strength in relation to morphological measures of
- 925 masculinity, fluctuating asymmetry and sexual behaviour in males and females. Sex
- 926 *Hormones. InTech*, 293-306. (2012).
- 927 77. C. R. von Rueden, A. V. Jaeggi, Men's status and reproductive success in 33 nonindustrial
- 928 societies: Effects of subsistence, marriage system, and reproductive strategy. *Proc. Natl.*
- 929 *Acad. Sci. U.S.A.* 113(39), 10824-10829 (2016).
- 930 78. Y. Xu, S. Norton, Q. Rahman, Early life conditions, reproductive and sexuality-related life
- history outcomes among human males: A systematic review and meta-analysis. *Evol.*
- 932 *Hum. Behav.* 39(1), 40-51 (2018).

- 933 79. L. Penke, J. B. Asendorpf, Beyond global sociosexual orientations: A more differentiated
 934 look at sociosexuality and its effects on courtship and romantic relationships. *J. Pers. Soc.*935 *Psychol.* 95(5), 1113-1135 (2008).
- 80. J. A. Simpson, S. W. Gangestad, Individual differences in sociosexuality: Evidence for
 convergent and discriminant validity. *J. Pers. Soc. Psychol.* 60(6), 870-883 (1991).
- 938 81. A. Alvergne, C. Faurie, M. Raymond, Variation in testosterone levels and male
- 939 reproductive effort: Insight from a polygynous human population. *Horm. Behav.* 56(5),
 940 491-497 (2009).
- 82. C. L. Apicella, Upper-body strength predicts hunting reputation and reproductive success
 in Hadza hunter–gatherers. *Evol. Hum. Behav.* 35(6), 508-518 (2014).
- 83. C. L. Apicella, D. R. Feinberg, F. W. Marlowe, Voice pitch predicts reproductive success
 in male hunter-gatherers. *Biol. Lett.*, 3(6), 682-684 (2007).
- 84. J. Aronoff, "Life history strategies, testosterone, and the anthropology of human
 development," University of Alabama, Tuscaloosa, Alabama (2017).
- 947 85. J. A. Atkinson, "Anthropometric correlates of reproductive success, facial configuration,
- risk taking and sexual behaviors among indigenous and Western populations: The role of
- hand-grip strength and wrist width," University at Albany, Albany, New York (2012).
- 950 86. J. Atkinson, R. N. Pipitone, A. Sorokowska, P. Sorokowski, M. Mberira, A. Bartels, G. G.
- 951 Gallup, Voice and handgrip strength predict reproductive success in a group of indigenous
- 952 African females. *PLoS One* 7(8), e41811 (2012).
- 87. A. F. Bogaert, W. A. Fisher, Predictors of university men's number of sexual partners. *J. Sex Res.* 32(2), 119-130 (1995).
- 88. A. Booth, D. R. Johnson, D. A. Granger, Testosterone and men's health. *J. Behav. Med.*22(1), 1-19 (1999).

- 957 89. L. G. Boothroyd, C. P. Cross, A. W. Gray, C. Coombes, K. Gregson-Curtis, Perceiving
- 958 the facial correlates of sociosexuality: Further evidence. *Pers. Individ.* 50(3), 422-425
 959 (2011).
- 960 90. L. G. Boothroyd, B. C. Jones, D. M. Burt, L. M. DeBruine, D. I. Perrett, Facial correlates
 961 of sociosexuality. *Evol. Hum. Behav.* 29(3), 211-218 (2008).
- 962 91. N. E. Charles, G. M. Alexander, The association between 2D:4D ratios and
- sociosexuality: A failure to replicate. *Arch. Sex. Behav.* 40(3), 587-595 (2011).
- 964 92. N. Chaudhary, et al. Polygyny without wealth: Popularity in gift games predicts polygyny

965 in BaYaka Pygmies. R. Soc. Open Sci. 2(5), 150054 (2015).

- 966 93. R. S. Edelstein, W. J. Chopik, E. L. Kean, Sociosexuality moderates the association
- between testosterone and relationship status in men and women. *Horm. Behav.* 60(3), 248255 (2011).
- 969 94. R. G. Falcon, Stay, stray or something in-between? A comment on Wlodarski et al. *Biol.*970 *Lett.* 12(5), 20151069 (2016).
- 971 95. D. Farrelly, R. Owens, H. R. Elliott, H. R. Walden, M. A. Wetherell, The effects of being
- 972 in a "new relationship" on levels of testosterone in men. *Evol. Psychol.* 13(1),
- 973 147470491501300116 (2015).
- 96. M. J. Frederick, "Effects of early developmental stress on adult physiology and behavior,"
 State University of New York at Albany, Albany, New York (2010).
- 976 97. D. A. Frederick, B. N. Jenkins, Height and body mass on the mating market: Associations
 977 with number of sex partners and extra-pair sex among heterosexual men and women aged
- 978 18–65. *Evol. Psychol.* 13(3) (2015).
- 979 98. A. C. Gallup, D. D. White, G. G. Gallup Jr, Handgrip strength predicts sexual behavior,
- body morphology, and aggression in male college students. *Evol. Hum. Behav.* 28(6), 423-
- 981 429 (2007).

- 982 99. J. E. Genovese, Physique correlates with reproductive success in an archival sample of
 983 delinquent youth. *Evol. Psychol.* 6(3) (2008).
- 984 100. L. T. Gettler, P. X. Kuo, S. Rosenbaum, J. L. Avila, T. W. McDade, C. W. Kuzawa,

985 Sociosexuality, testosterone, and life history status: Prospective associations and

- 986 longitudinal changes among men in Cebu, Philippines. *Evol. Hum. Behav.* 40(2), 249-258
- 987 (2019).
- 101. T. E. Gildner, "Life history tradeoffs between testosterone and immune function
 among Shuar forager-horticulturalists of Amazonian Ecuador," University of Oregon,
 Eugene, Oregon (2018).
- 991 102. J. Gómez-Valdés, et al, Lack of support for the association between facial shape and
 992 aggression: A reappraisal based on a worldwide population genetics perspective. *PLoS*993 *One* 8(1), e52317 (2013).
- 103. E. M. Hartl, E. P. Monnelly, R. D. Elderkin, *Physique and delinquent behavior: A thirty year follow-up of William H. Sheldon's Varieties of Delinquent Youth* (Academic
 Press, 2013).
- 997 104. S. Hoppler, A. Walther, P. La Marca-Ghaemmaghami, U. Ehlert, Lower birthweight
 998 and left-/mixed-handedness are associated with intensified age-related sex steroid decline
- in men. Findings from the Men's Health 40+ Study. *Andrology* 6(6), 896-902 (2018).
- 1000 105. J. Hönekopp, M. Voracek, J. T. Manning, 2nd to 4th digit ratio (2D:4D) and number
- 1001 of sex partners: Evidence for effects of prenatal testosterone in men.
- 1002 *Psychoneuroendocrinology* 31(1), 30-37 (2006).
- 1003 106. J. Hönekopp, U. Rudolph, L. Beier, A. Liebert, C. Müller, Physical attractiveness of
 1004 face and body as indicators of physical fitness in men. *Evol. Hum. Behav.* 28(2), 106-111
 1005 (2007).

- 1006 107. S. Kirchengast, Differential reproductive success and body size in! Kung San people
 1007 from northern Namibia. *Coll. Antropol.* 24(1), 121-132 (2000).
- 1008 108. S. Kirchengast, E. M. Winkler, Differential reproductive success and body dimensions
 in Kavango males from urban and rural areas in northern Namibia. *Hum. Biol.* 291-309
- 1010 (1995).
- 1011 109. C. Klimas, U. Ehlert, T. J. Lacker, P. Waldvogel, A. Walther, Higher testosterone
 1012 levels are associated with unfaithful behavior in men. *Biol. Psychol.* 146, 107730 (2019).
- 1013 110. M. Klimek, A. Galbarczyk, I. Nenko, L. C. Alvarado, G. Jasienska, Digit ratio
- 1014 (2D:4D) as an indicator of body size, testosterone concentration and number of children in
 1015 human males. *Ann. Hum. Biol.* 41(6), 518-523 (2014).
- 1016 111. T. L. Kordsmeyer, L. Penke, The association of three indicators of developmental
 1017 instability with mating success in humans. *Evol. Hum. Behav.* 38(6), 704-713 (2017).
- 1018 112. M. Krzyżanowska, C. N. Mascie-Taylor, J. C. Thalabard, Is human mating for height
- associated with fertility? Results from a British national cohort study. *Am. J. Hum. Biol.*27(4), 553-563 (2015).
- 1021 113. R. Kurzban, J. Weeden, HurryDate: Mate preferences in action. *Evol. Hum. Behav.*1022 26(3), 227-244 (2005).
- 1023 114. B. B. Little, R. M. Malina, P. H. Buschang, L. R. Little, Natural selection is not related
 1024 to reduced body size in a rural subsistence agricultural community in southern Mexico.
- 1025 *Hum. Biol.* 61(2), 287-296 (1989).
- 1026 115. J. Loehr, R. B. O'Hara, Facial morphology predicts male fitness and rank but not
 1027 survival in Second World War Finnish soldiers. *Biol. Lett.* 9(4), 20130049 (2013).
- 1028 116. D. P. Longman, M. K. Surbey, J. T. Stock, J. C. Wells, Tandem androgenic and
- 1029 psychological shifts in male reproductive effort following a manipulated "win" or "loss"
- 1030 in a sporting competition. *Hum. Nat.* 29(3), 283-310 (2018).

- 1031 117. V. X. Luevano, K. Merkling, J. J. Sanchez-Pinelo, M. Shakeshaft, M. L. Velazquez,
- 1032 "Facial width-to-height ratio, psychopathy, sociosexual orientation, and accepting
- 1033 detection-risk in extra-pair relationships." Poster presented at the annual meeting of the
- 1034 Western Psychological Association, Portland, OR. (2018, April).
- 1035 118. D. Maestripieri, A. Klimczuk, D. Traficonte, M. C. Wilson, Ethnicity-related variation
 1036 in sexual promiscuity, relationship status, and testosterone levels in men. *Evol. Behav. Sci.*1037 8(2), 96 (2014).
- 1038 119. J. T. Manning, B. Fink, Digit ratio (2D:4D), dominance, reproductive success,
- asymmetry, and sociosexuality in the BBC Internet Study. *Am. J. Hum. Biol.* 20(4), 451461 (2008).
- 1041 120. J. T. Manning, P. Henzi, P. Venkatramana, S. Martin, D. Singh, Second to fourth digit
 ratio: Ethnic differences and family size in English, Indian and South African populations.
- 1043 Ann. Hum. Biol. 30(5), 579-588 (2003).
- 1044 121. M. Marczak, M. Misiak, A. Sorokowska, P. Sorokowski, No sex difference in digit
- ratios (2D:4D) in the traditional Yali of Papua and its meaning for the previous
- 1046 hypotheses on the inter-populational variability in 2D:4D. Am. J. Hum. Biol. 30(2) (2018).
- 1047 122. M. McIntyre, S. W. Gangestad, P. B. Gray, J. F. Chapman, T. C. Burnham, M. T.
- 1048 O'Rourke, R. Thornhill, Romantic involvement often reduces men's testosterone levels--
- 1049 but not always: The moderating role of extrapair sexual interest. J. Pers. Soc. Psychol.
- 1050 91(4), 642 (2006).
- 1051 123. J. Međedović, T. Bulut, A life-history perspective on body mass: Exploring the
 1052 interplay between harsh environment, body mass, and mating success. *Evol. Behav. Sci.*1053 13(1), 84 (2019).
- 1054 124. M. A. Mosing, K. J. Verweij, G. Madison, N. L. Pedersen, B. P. Zietsch, F. Ullén, Did
 1055 sexual selection shape human music? Testing predictions from the sexual selection

- 1056 hypothesis of music evolution using a large genetically informative sample of over 10,000
- 1057 twins. Evol. Hum. Behav. 36(5), 359-366 (2015).
- 1058 125. U. Muller, A. Mazur, Facial dominance in Homo sapiens as honest signaling of male
 1059 quality. *Behav. Ecol.* 8(5), 569-579 (1997).
- 1060 126. N. J. Nagelkerke, R. M. Bernsen, S. K. Sgaier, P. Jha, Body mass index, sexual
- behaviour, and sexually transmitted infections: An analysis using the NHANES 1999–
 2000 data. *BMC Public Health* 6(1), 199 (2006).
- 1063 127. D. Nettle, Height and reproductive success in a cohort of British men. *Hum. Nat.*1064 13(4), 473-491 (2002).
- 1065 128. B. Pawlowski, R. Dunbar, A. Lipowicz, Tall men have more reproductive success.
 1066 *Nature* 403 (6766), 156 (2000).
- 1067 129. T. V. Pollet, L. van der Meij, K. D. Cobey, A. P. Buunk, Testosterone levels and their
 associations with lifetime number of opposite sex partners and remarriage in a large
- sample of American elderly men and women. *Horm. Behav.* 60(1), 72-77 (2011).
- 1070 130. P. Polo, J. A. Muñoz-Reyes, M. Pita, T. K. Shackelford, B. Fink, Testosterone-
- dependent facial and body traits predict men's sociosexual attitudes and behaviors. *Am. J. Hum. Biol.* 31(3), e23235 (2019).
- 1073 131. M. E. Price, N. Pound, J. Dunn, S. Hopkins, J. Kang, Body shape preferences:
- 1074 Associations with rater body shape and sociosexuality. *PLoS One* 8(1) (2013).
- 1075 132. P. Prokop, P. Fedor, Physical attractiveness influences reproductive success of modern
 1076 men. *J Ethol* 29(3), 453-458 (2011).
- 1077 133. P. Prokop, P. Fedor, Associations between body morphology, mating success and
- 1078 mate preferences among Slovak males and females. *Anthropol. Anz.* 70(2), 121-135
- 1079 (2013).

1080	134. D. A. Puts, S. J. Gaulin, K. Verdolini, Dominance and the evolution of sexual
1081	dimorphism in human voice pitch. Evol. Hum. Behav. 27(4), 283-296 (2006).
1082	135. D. A. Puts, L. E. Pope, A. K. Hill, R. A., Cárdenas, L. L. Welling, J. R. Wheatley, S.
1083	M. Breedlove, Fulfilling desire: Evidence for negative feedback between men's
1084	testosterone, sociosexual psychology, and sexual partner number. Horm. Behav. 70, 14-21
1085	(2015).
1086	136. D. A. Putz, S. J. Gaulin, R. J. Sporter, D. H. McBurney, Sex hormones and finger
1087	length: What does 2D:4D indicate?. Evol. Hum. Behav. 25(3), 182-199 (2004).
1088	137. Q. Rahman, M. Korhonen, A. Aslam, Sexually dimorphic 2D:4D ratio, height, weight,
1089	and their relation to number of sexual partners. Pers. Individ. 39(1), 83-92 (2005).
1090	138. K. A. Rosenfield, A. Sorokowska, P. Sorokowski, D. A. Puts, Sexual selection for low
1091	male voice pitch among Amazonian forager-horticulturists. Evol. Hum. Behav. 41(1), 3-11

- 1092 (2020).
- 1093 139. S. Schwarz, M. Mustafić, M. Hassebrauck, J. Jörg, Short-and long-term relationship
 1094 orientation and 2D:4D finger-length ratio. *Arch. Sex. Behav.* 40(3), 565-574 (2011).
- 1095 140. E. C. Scott, C. J. Bajema, Height, weight and fertility among the participants of the
 1096 third Harvard growth study. *Hum. Biol.* 54(3), 501-516 (1982).
- 1097 141. M. L. Shoup, G. G. Gallup, Men's faces convey information about their bodies and
 1098 their behavior: What you see is what you get. *Evol. Psychol.* 6(3), 147470490800600311
 1099 (2008).
- 142. K. Sim, W. Y. Chun, The relationships of assertiveness and responsiveness to sexual
 behavior. *Psychologia* 59(1), 50-69 (2016).
- 1102 143. Z. L. Simmons, J. R. Roney, Variation in CAG repeat length of the androgen receptor
- 1103 gene predicts variables associated with intrasexual competitiveness in human males.
- 1104 *Horm. Behav.* 60(3), 306-312 (2011).

144. K. M. Smith, Y. M. Olkhov, D. A. Puts, C. L. Apicella, Hadza men with lower voice
pitch have a better hunting reputation. *Evol. Psychol.* 15(4), 1474704917740466 (2017).

1107 145. M. Sneade, A. Furnham, Hand grip strength and self-perceptions of physical

- attractiveness and psychological well-being. *Evol. Psychol. Sci.* 2(2), 123-128 (2016).
- 1109 146. P. Sorokowski, A. Sorokowska, D. P. Danel, Why pigs are important in Papua?
- Wealth, height and reproductive success among the Yali tribe of West Papua. *Econ. Hum. Biol.* 11(3), 382-390 (2013).
- 1112 147. E. T. Steiner, "Testosterone and vasopressin in men's reproductive behavior,"
- 1113 University of Oregon, Eugene, Oregon (2011).
- 1114 148. J. Stern, K. Karastoyanova, M. Kandrik, J. Torrance, A. C. Hahn, I. Holzleitner,... B.
- 1115 C. Jones, Are sexual desire and sociosexual orientation related to men's salivary steroid
 1116 hormones?. *Adapt. Hum. Behav. Physiol.* 6(4) 447-466 (2020).
- 1117 149. H. L. Strong, "Association of prenatal androgen exposure with attention and
- 1118 sociosexual orientation," California State University, Long Beach, California (2014).
- 1119 150. H. Strong, V. X. Luevano, "Prenatal androgen exposure predicts relationship-type
- 1120 preference but not experience," Poster presented at the annual meeting of the Association
- 1121 for Psychological Science, San Diego, CA. (2014, May).
- 1122 151. S. V. Subramanian, L. K. Ackerson, G. D. Smith, N. A. John, Association of maternal
- height with child mortality, anthropometric failure, and anemia in India. *JAMA* 301(16)
- 1124 1691-1701 (2009).
- 1125 152. A. Suire, M. Raymond, M. Barkat-Defradas, Human vocal behavior within
- 1126 competitive and courtship contexts and its relation to mating success. *Evol. Hum. Behav.*
- 1127 39(6), 684-691 (2018).
- 1128 153. H. L., Tao, C. C. Yin, Height, partners and offspring: Evidence from Taiwan. J.
- 1129 *Biosoc. Sci.* 48(5) 593-615 (2016).

- 1130 154. S. M. van Anders, L. D. Hamilton, N. V. Watson, Multiple partners are associated
 1131 with higher testosterone in North American men and women. *Horm. Behav.* 51(3), 4541132 459 (2007).
- 1133 155. M. A. C. Varella, J. V. Valentova, K. J. Pereira, V. S. R. Bussab, Promiscuity is
- related to masculine and feminine body traits in both men and women: Evidence from
 Brazilian and Czech samples. *Behav. Process.* 109, 34-39 (2014).
- 1136 156. C. von Rueden, M. Gurven, H. Kaplan, Why do men seek status? Fitness payoffs to
 1137 dominance and prestige. *Proc. R. Soc. B.* 278(1715), 2223-2232 (2011).
- 1138 157. M. Voracek, U. Pum, S. G. Dressler, Investigating digit ratio (2D:4D) in a highly
- male-dominated occupation: The case of firefighters. *Scand. J. Psychol.* 51(2), 146-156
 (2010).
- 1141 158. A. Walther, M. Philipp, N. Lozza, U. Ehlert, The rate of change in declining steroid
 1142 hormones: a new parameter of healthy aging in men?. *Oncotarget* 7(38), 60844 (2016).
- 1143 159. A. Walther, M. Philipp, N. Lozza, U. Ehlert, Emotional support, depressive symptoms,
- and age-related alterations in male body composition: Cross-sectional findings from the
- 1145 men's health 40+ study. *Front. Psychol.* 8, 1075 (2017).
- 1146 160. A. Walther, F. Mahler, R. Debelak, U. Ehlert, Psychobiological protective factors
- modifying the association between age and sexual health in men: Findings from the men's
 health 40+ study. *Am. J. Men's Health* 11(3), 737-747 (2017).
- 1149 161. A. Walther, P. Waldvogel, E. Noser, J. Ruppen, U. Ehlert, Emotions and steroid
 1150 secretion in aging men: A multi—study report. *Front. Psychol.* 8, 1722 (2017).
- 1151 162. D. Waynforth, Fluctuating asymmetry and human male life-history traits in rural
 1152 Belize. *Proc. R. Soc. B* 265(1405), 1497-1501 (1998).
- 1153 163. E. M. Winkler, S. Kirchengast, Body dimensions and differential fertility in! Kung
- 1154 San males from Namibia. *Am. J. Hum. Biol.* 6(2), 203-213 (1994).

- 1155 164. W. Viechtbauer, Conducting meta-analyses in R with the metafor package. J. Stat.
 1156 Softw. 36(3), 1–48 (2010).
- 1157 165. R Development Core Team (2019) R: A Language and Environment for Statistical
 1158 Computing. R Foundation for Statistical Computing, Vienna.
- 1159 166. J. D. Storey, A direct approach to false discovery rates, *J. R. Stat. Soc.* (64) 479–498
 1160 (2002).
- 1161 167. The World Bank. (2018). Fertility rate, total (births per woman).
- 1162 <u>https://data.worldbank.org/indicator/SP.DYN.TFRT.IN</u>
- 1163 168. D. C. Funder, D. J. Ozer, Evaluating effect size in psychological research: Sense and
 1164 nonsense. *AMPPS* 2(2), 156-168 (2019).
- 1165 169. A. Gat, Proving communal warfare among hunter-gatherers: The quasi-rousseauan
 1166 error. *Evol. Anthropol.* 24(3), 111-126 (2015).
- 1167 170. L. Glowacki, R. Wrangham, Warfare and reproductive success in a tribal population.
 1168 *Proc. Natl. Acad. Sci. U.S.A.* 112(2), 348-353 (2015).
- 1169 171. J. H. Manson, R. W. Wrangham, J. L. Boone, B. Chapais, R. I. M. Dunbar, R. Ember,
- 1170 W. Irons, L. F. Marchant, W. C. McGrew, T. Nishida, Paterson, J. D. Smith, E. A.
- 1171 Stanford, C. B. C. M. Worthman, Intergroup aggression in chimpanzees and humans [and
- 1172 comments and replies]. *Curr. Anthropol.* 32(4), 369-390 (1991).
- 1173 172. J. M. Plavcan, Body size, size variation, and sexual size dimorphism in early Homo.
- 1174 *Curr. Anthropol.* 53(S6), S409-S423 (2012).
- 1175 173. N. A. Chagnon, Life histories, blood revenge, and warfare in a tribal population.
- 1176 Science 239(4843), 985-992 (1988).
- 1177 174. V. S. Johnston, R. Hagel, M. Franklin, B. Fink, K. Grammer, Male facial
- 1178 attractiveness: Evidence for hormone-mediated adaptive design. *Evol. Hum. Behav.* 22(4),
- 1179 251-267 (2001).

- 1180 175. R. G. Bribiescas, Testosterone levels among Aché hunter-gatherer men. *Hum. Nat.*1181 7(2), 163-188 (1996).
- 1182 176. J. Archer, Testosterone and human aggression: An evaluation of the challenge
 1183 hypothesis. *Neurosci Biobehav Rev* 30(3), 319-345 (2006).
- 1184 177. S. A. Holmboe, L. Priskorn, N. Jørgensen, N. E. Skakkebaek, A. Linneberg, A. Juul,
- 1185 A.-M. Andersson, Influence of marital status on testosterone levels–A ten year follow-up
- 1186 of 1113 men. *Psychoneuroendocrinology* 80, 155-161 (2017).
- 1187 178. L. T. Gettler, T. W. McDade, A. B. Feranil, C. W. Kuzawa, Longitudinal evidence
- that fatherhood decreases testosterone in human males. *Proc. Natl. Acad. Sci. U.S.A.*
- 1189 108(39), 16194-16199 (2011).
- 1190

1191 Supplementary References

- 1192 A. P. Clark, Self-perceived attractiveness and masculinization predict women's
- 1193 sociosexuality. *Evol. Hum. Behav.* 25(2), 113-124 (2004).
- 1194 C. R. Hodges-Simeon, S. J. Gaulin, D. A. Puts, Voice correlates of mating success in men:
- examining "contests" versus "mate choice" modes of sexual selection. *Arch. Sex. Behav.* 40(3), 551-557 (2011).
- 1197 J. Hönekopp, M. Voracek, J. T. Manning, 2nd to 4th digit ratio (2D:4D) and number of sex
- 1198partners: Evidence for effects of prenatal testosterone in men.
- 1199 *Psychoneuroendocrinology* 31(1), 30-37 (2006).
- 1200 M. Kandrik, A. C. Hahn, C. Han, J. Wincenciak, C. I. Fisher, L. M. DeBruine, B. C. Jones,
- 1201 Does the interaction between cortisol and testosterone predict men's facial
- 1202 attractiveness?. Adapt. Hum. Behav. Physiol. 3(4), 275-281 (2017).
- 1203 M. G. Kendall, *Rank correlation methods (4th ed.)*. (Charles Griffin & Co., 1970).

1204	R. A. Peterson, S. P. Brown, On the use of beta coefficients in meta-analysis. Adapt. Hum.
1205	Behav. Physiol. 90(1), 175 (2005).
1206	K. Sim, The relationship between sex-typical body shape and quality indicators. JSEC 7(2),
1207	97 (2013).
1208	
1209	LIST OF FIGURES, TABLES AND SUPPLEMENTARY FILES (figures submitted as
1210	separate files; figure captions below)
1211	
1212	Figure 1
1213	Figure 1. All studies included in the meta-analysis.
1214	
1215	Figure 2
1216	Figure 2. Overall analysis structure.
1217	
1218	Figure 3
1219 1220 1221	<i>Figure 3</i> . Forest plot of the association between body masculinity and the mating domain. Effect sizes are shown as Z-transformed r , with 95% confidence intervals in brackets. The width of the diamond corresponds to the confidence interval for the overall effect.
1222	
1223	Figure 4
1224 1225 1226 1227	<i>Figure 4</i> . Forest plot of the association between voice pitch and the mating domain. Effect sizes are shown as <i>Z</i> -transformed <i>r</i> , with 95% confidence intervals in brackets. The width of the diamond corresponds to the confidence interval for the overall effect.
1228	Figure 5
1229 1230 1231	<i>Figure 5.</i> Forest plot of the association between testosterone levels and the mating domain. Effect sizes are shown as Z-transformed r , with 95% confidence intervals in brackets. The width of the diamond corresponds to the confidence interval for the overall effect.

1233 Figure 6

Figure 6. Forest plot of the association between height and the mating domain. Effect sizes
are shown as *Z*-transformed *r*, with 95% confidence intervals in brackets. The width of the
diamond corresponds to the confidence interval for the overall effect.

- 1237
- 1238 Figure 7

1239 *Figure 7.* Forest plot of the association between body masculinity and the reproductive

- 1240 domain. Effect sizes are shown as Z-transformed *r*, with 95% confidence intervals in brackets.
- 1241 The width of the diamond corresponds to the confidence interval for the overall effect.
- 1242

Table 1

Masculine traits predicting mating: main analyses and subgroup analyses of mating attitudes vs mating behaviors and low vs high fertility samples. Pearson's r (95% CI); p value for meta-analytic effect, q-value (correcting for multiple comparisons); number of observations (k), samples (s), and unique participants (n); test for heterogeneity (Q), p value for heterogeneity. Statistically significant meta-analytic associations are bolded if still significant after controlling for multiple comparisons.

			Ma	ting		
Outcome: Sample	Facial masculinity	Body masculinity	2D:4D	Voice pitch	Height	T levels
Mating domain: All samples	r = .080 (-0.003, 0.164), p = .060, q = .117 k = 30, s = 11, n = 948 Q(df = 29) = 54.834, p = .003	r = .133 (0.091, 0.176), p < .001, q = .001 k = 121, s = 32, n = 7939 Q(df = 120) = 297.472, p < .001	r = .034 (0.000, 0.069), p = .049, q = .102 k = 84, s = 23, n = 66807 Q(df = 83) = 101.994, p = .077	r = .132 (0.061, 0.204), p < .001, q = .002 k = 8, s = 5, n = 443 Q(df = 7) = 2.334, p = .939	r = .057 (0.027, 0.087), p < .001, q = .002 k = 62, s = 25, n = 43686 Q(df = 61) = 263.247, p < .001	r = .093 (0.066, 0.121), p < .001, q = .001 k = 66, s = 21, n = 7083 Q(df = 65) = 66.090, p = .439
Mating attitudes: All samples	r = .095 (-0.072, 0.263), p = .263, q = .304 k = 5, s = 4, n = 407 Q(df = 4) = 8.684, p = .070	r = .078 (0.002, 0.155), p = .045, q = .098 k = 20, s = 9, n = 922 Q(df = 19) = 17.606, p = .549	r = .035 (-0.061, 0.132), p = .474, q = .385 k = 19, s = 7, n = 504 Q(df = 18) = 24.141, p = .151	<i>s</i> = 0	r = .028 (-0.013, 0.068), p = .179, q = .253 k = 9, s = 6, n = 4232 Q(df = 8) = 5.137, p = .743	<i>r</i> = .099 (0.026, 0.173), <i>p</i> = .008, <i>q</i> = .032 <i>k</i> = 21, <i>s</i> = 11, <i>n</i> = 1039 Q(df = 20) = 25.379, <i>p</i> = .187
Mating behaviors: All samples	r = .025 (-0.059, 0.109), p = .554, q = .424 k = 22, s = 8, n = 755 Q(df = 21) = 37.044, p = .017	r = .142 (0.099, 0.187), p < .001, q = .001 k = 91, s = 31, n = 7738 Q(df = 90) = 267.876, p < .001	r = .038 (-0.002, 0.078), p = .061, q = .117 k = 51, s = 19, n = 1607 Q(df = 50) = 64.049, p = .087	r = .124 (0.043, 0.206), p = .003, q = .016 k = 7, s = 5, n = 443 Q(df = 6) = 2.162, p = .904	r = .054 (0.021, 0.087), p = .001, q = .008 k = 48, s = 24, n = 42179 Q(df = 47) = 247.032, p < .001	r = .084 (0.058, 0.110), p < .001, q = .001 k = 32, s = 17, n = 6765 Q(df = 31) = 28.558, p = .592
Mating domain: Low fert. samples	r = .089 (-0.001, 0.179), p = .053, q = .109 k = 28, s = 10, n = 913	r = .135 (0.091, 0.180), p < .001, q = .001 k = 117, s = 28, n = 7572	r = .038 (0.002, 0.073), p = .037, q = .086 k = 82, s = 22, n = 66751	r = .129 (0.055, 0.204), p < .001, q = .005 k = 7, s = 4, n = 388	r = .055 (0.024, 0.086), p < .001, q = .004 k = 58, s = 21, n = 43310	r = .099 (0.069, 0.129), p < .001, q = .001 k = 58, s = 20, n = 6795

	Q(df = 27) = 54.287, p = .001	Q(df = 116) = 289.080, p < .001	Q(df = 81) = 101.369, p = .063	Q(df = 6) = 2.234, p = .897	Q(df = 57) = 259.576, p < .001	Q(df = 57) = 61.443, p = .320
Mating attitudes: Low fert. samples	r = .095 (-0.072, 0.262), p = .263, q = .304 k = 5, s = 4, n = 407 Q(df = 4) = 8.684, p = .070	r = .078 (0.002, 0.155), p = .045, q = .098 k = 20, s = 9, n = 922 Q(df = 19) = 17.606, p = .549	r = .035 (-0.061, 0.132), p = .474, q = .385 k = 19, s = 7, n = 504 Q(df = 18) = 24.141, p = .151	<i>s</i> = 0	r = .028 (-0.013, 0.068), p = .179, q = .253 k = 9, s = 6, n = 4232 Q(df = 8) = 5.137, p = .743	r = .108 (0.021, 0.195), p = .015, q = .047 k = 17, s = 10, n = 751 Q(df = 16) = 20.017, p = .220
Mating behaviors: Low fert. samples	r = .028 (-0.063, 0.119), p = .543, q = .420 k = 20, s = 7, n = 720	<i>r</i> = .145 (0.100, 0.193), <i>p</i> < .001, <i>q</i> = .001 <i>k</i> = 87, <i>s</i> = 27, <i>n</i> = 7371	r = .042 (0.001, 0.083), p = .045, q = .098 k = 49, s = 19, n = 1551	<i>r</i> = .119 (0.034, 0.205), <i>p</i> = .006, <i>q</i> = .025 <i>k</i> = 6, <i>s</i> = 4, <i>n</i> = 388	r = .051 (0.017, 0.086), p = .004, q = .019 k = 44, s = 20, n = 41803	r = .088 (0.058, 0.119), p < .001, q = .001 k = 30, s = 16, n = 6477
	Q(df = 19) = 36.610, p = .009	Q(df = 86) = 259.448, p < .001	Q(df = 48) = 62.941, p = .073	Q(df = 5) = 2.017, p = .847	Q(df = 43) = 243.392, p < .001	Q(df = 29) = 27.793, p = .529
Mating domain: High fert. samples	<i>s</i> = 1	r = .105 (-0.069, 0.280), p = .235, q = .285 k = 4, s = 4, n = 367 Q(df = 3) = 7.282, p = .063	<i>s</i> = 1	<i>s</i> = 1	r = .089 (-0.016, 0.193), p = .096, q = .157 k = 4, s = 4, n = 376 Q(df = 3) = 3.388, p = .336	<i>s</i> = 1
Mating attitudes: High fert. samples	s = 0	s = 0	s = 0	s = 0	s = 0	<i>s</i> = 1
Mating behaviors: High fert. samples	<i>s</i> = 1	r = .105 (-0.069, 0.280), p = .235, q = .285 k = 4, s = 4, n = 367 Q(df = 3) = 7.282, p = .063	<i>s</i> = 1	<i>s</i> = 1	r = .089 (-0.016, 0.193), p = .096, q = .157 k = 4, s = 4, n = 376 Q(df = 3) = 3.388, p = .336	<i>s</i> = 1

Note. Fert. = fertility; k = number of observations; n = number of unique participants; Q = Cochran's Q test of heterogeneity; q = q-value; s = number of samples; T = testosterone. 1243

1244

Table 2

Masculine traits predicting reproduction: main analyses and subgroup analyses of mating attitudes vs mating behaviors and low vs high fertility samples. Pearson's r (95% CI); p value for meta-analytic effect, q-value (correcting for multiple comparisons); number of observations (k), samples (s), and unique participants (n); test for heterogeneity (Q), p value for heterogeneity. Statistically significant meta-analytic associations are bolded if still significant after controlling for multiple comparisons.

			Repro	duction		
Outcome: Sample	Facial masculinity	Body masculinity	2D:4D	Voice pitch	Height	T levels
Reproductive domain: All samples	r = .099 (-0.012, 0.211), p = .081, q = .140 k = 5, s = 5, n = 1411 Q(df = 4) = 8.799, p = .066	r = .143 (0.076, 0.212), p < .001, q = .001 k = 14, s = 8, n = 897 Q(df = 13) = 16.356, p = .230	r = .074 (-0.006, 0.154), p = .070, q = .131 k = 19, s = 10, n = 84558 Q(df = 18) = 31.704, p = .024	r = .136 (-0.053, 0.328), p = .158, q = .228 k = 5, s = 3, n = 143 Q(df = 4) = 5.378, p = .251	r = .006 (-0.049, 0.062), p = .819, q = .491 k = 35, s = 25, n = 22326 Q(df = 34) = 433.359, p < .001	r = .039 (-0.067, 0.145), p = .474, q = .385 k = 3, s = 3, n = 351 Q(df = 2) = 0.387, p = .824
Fertility: All samples	r = .003 (-0.253, 0.260), p = .980, q = .543 k = 3, s = 3, n = 437 Q(df = 2) = 5.416, p = .067	r = .130 (0.060, 0.201), p < .001, q = .002 k = 8, s = 6, n = 813 Q(df = 7) = 4.840, p = .679	r = .032 (-0.065, 0.130), p = .514, q = .406 k = 13, s = 5, n = 84128 Q(df = 12) = 17.757, p = .123	<i>s</i> = 2	r = .011 (-0.039, 0.062), p = .660, q = .451 k = 26, s = 23, n = 22242 Q(df = 25) = 400.038, p < .001	<i>s</i> = 2
RS: All samples	<i>s</i> = 2	r = .192 (-0.052, 0.441), p = .122, q = .189 k = 6, s = 4, n = 205 Q(df = 5) = 11.344, p = .045	r = .174 (0.085, 0.267), p < .001, q = .002 k = 6, s = 5, n = 430 Q(df = 5) = 0.976, p = .965	<i>s</i> = 2	r = -0.044 (-0.201, 0.113), p = .584, q = .430 k = 9, s = 9, n = 603 Q(df = 8) = 33.311, p < .001	<i>s</i> = 1
Reproductive domain: Low fert. samples	s = 0	<i>s</i> = 1	<i>r</i> = .083 (-0.023, 0.190), <i>p</i> = .126, <i>q</i> = .191	s = 0	<i>r</i> = -0.037 (-0.112, 0.038), <i>p</i> = .337, <i>q</i> = .347	<i>s</i> = 2

			k = 8, s = 4, n = 84034 Q(df = 7) = 13.988, p = .051		k = 8, s = 8, n = 17135 Q(df = 7) = 244.970, p < .001	
Fertility: Low fert. samples	s = 0	<i>s</i> = 1	<i>r</i> = .052 (-0.065, 0.169), <i>p</i> = .386, <i>q</i> = .369	s = 0	<i>r</i> = -0.037 (-0.112, 0.038), <i>p</i> = .337, <i>q</i> = .347	<i>s</i> = 2
			<i>k</i> = 7, <i>s</i> = 3, <i>n</i> = 83845		k = 8, s = 8, n = 17135	
			Q(df = 6) = 8.335, p = .215		Q(df = 7) = 244.970, <i>p</i> < .001	
RS: Low fert. samples	s = 0	s = 0	<i>s</i> = 1	s = 0	s = 0	s = 0
Reproductive domain: High fert. samples	r = .099 (-0.012, 0.211), p = .081, q = .140 k = 5, s = 5, n = 1411 Q(df = 4) = 8.799, p = .066	r = .163 (0.104, 0.225), p < .001, q = .001 k = 13, s = 7, n = 626 Q(df = 12) = 12.347, p = .418	r = .083 (-0.039, 0.205), p = .184, q = .257 k = 11, s = 6, n = 524 Q(df = 10) = 12.595, p = .247	r = .136 (-0.053, 0.327), p = .158, q = .228 k = 5, s = 3, n = 143 Q(df = 4) = 5.378, p = .251	r = .034 (-0.041, 0.109), p = .377, q = .367 k = 27, s = 17, n = 5191 Q(df = 26) = 70.216, p < .001	<i>s</i> = 1
Fertility: High fert. samples	r = .003 (-0.253, 0.260), p = .980, q = .543 k = 3, s = 3, n = 437 Q(df = 2) = 5.416, p = .067	r = .165 (0.095, 0.237), p < .001, q = .001 k = 7, s = 5, n = 542 Q(df = 6) = 0.988, p = .986	<i>s</i> = 2	<i>s</i> = 2	r = .059 (0.007, 0.111), p = .025, q = .068 k = 18, s = 15, n = 5107 Q(df = 17) = 26.458, p = .067	<i>s</i> = 0
RS: High fert. samples	<i>s</i> = 2	<i>r</i> = .192 (-0.052, 0.441), <i>p</i> = .122, <i>q</i> = .189	$r = .170 \ (0.053, \ 0.291),$ $p = .005, \ q = .022$	<i>s</i> = 2	<i>r</i> =044 (-0.201, 0.113), <i>p</i> = .584, <i>q</i> = .430	<i>s</i> = 1
		k = 6, s = 4, n = 205 Q(df = 5) = 11.344, p = .045	k = 5, s = 4, n = 241 Q(df = 4) = 0.965, p = .915		k = 9, s = 9, n = 603 Q(df = 8) = 33.311, p < .001	

 $\frac{p - .045}{Note. \text{ fert.} = \text{ fertility; } k = \text{ number of observations; } n = \text{ number of unique participants; } Q = \text{ Cochran's } Q \text{ test of heterogeneity; } q = q \text{-value; } \text{RS} = \text{reproductive success; } s = \text{ number of samples; } T = \text{testosterone.}}$

Table 3

Overview of moderation analyses for the mating vs reproductive domains. Significant associations are indicated by + and - signs, showing the direction of the moderator relative to the reference category (stated first in the moderator column); crosses indicate no significant moderation; and 'na' indicates that power was too low to run that specific analysis. Only associations that remained significant after controlling for multiple comparisons are indicated here.

Note that this table only shows general moderators shared by all masculine traits; for trait-specific moderation analyses, see Supplementary Files 4. Likewise, for moderation analyses of the two mating domain measures attitudes and behaviors, and the two reproductive domain measures fertility and reproductive success, we also refer to Supplementary Files 4.

1	Fac	ial	Bo	dy	2D:	4D	Vo	ice	Hei	ght	T le	vels
	ma	sc.	masc.				pitch		6			
Moderator	MA	RE	MA	RE	MA	RE	MĂ	RE	MA	RE	MA	RE
	Т	Р	Т	Р	Т	Р	Т	Р	Т	Р	Т	Р
Mating vs												
reproductive	,))	C)	C)	C .	,	
domain												
Mating attitudes		na	V	na	N	na	na	na	×	na		na
vs behaviors	X	na	X	ma	X	ma	na	na	X	na	X	na
Fertility vs	na	na	na		na		na	na	na		na	na
reproductive	na	па	па	X	па	X	па	па	па	X	na	па
success												
Low vs high	na	na	~	na	na	~	na	na	~	~	na	na
fertility sample	na	na	X	na	na	X	na	na	X	X	na	na
Low fertility:	na	na	10	no	5	no	no	no	5	na	10	na
student	na	па	X	na	X	na	па	па	X	па	X	па
vs non-student												
sample												
High fertility:	na	na	na	na	na	~	na	na	na	~	na	na
traditional vs	na	na	na	na	na	X	na	na	na	X	na	na
industrialized												
sample												
Predominantly												
white vs	X	na	X	na	_	X	na	na	X	X	X	na
mixed/other/unkno												
wn ethnicity												
sample												
Monogamous vs												
non-monogamous	na	na	X	na	na	X	na	na	na	X	na	na
marriage system												
Published vs	10	na	10	10	5	10	no	no	10	10	10	na
non-published	X	na	X	X	X	X	na	na	X	X	X	na
results												
Peer reviewed vs	no	no	10	n 0	50	n 0	no	no	10	no	n 0	n 0
not peer reviewed	na	na	X	na	X	na	na	na	X	па	na	na
study												

Heterosexual vs												
gay/mixed/unknow	X	na	X	na	X	X	na	na	X	+	_	na
n sample												
Non-normality-												
transformed vs	na	na	X	X	+	X	na	na	X	X	+	na
transformed					1.1							
variables												
Non-converted vs	na	na	~		na	na	na	na	~	~	~	na
converted effect	na	na	~	+	nu	na	na	nu	~	~	X	iiu
sizes												
Age controlled for	~	na		na	×	×	na	na	×	~	~	na
vs not controlled	~	na	+							X	X	na
for												
Inclusion of non-	na	na	na	~	na	na	na	na	na	~	~	na
relevant control	m	na	na	X	m	ma	na	m	ma	X	X	ma
variables vs not												

Note. Masc = masculinity; MAT = mating; REP = reproduction; T = testosterone.

- 1248 Supplementary File 11249 *Effect size conversion formulas*
- 1250
- 1251 Supplementary File 2A
- 1252 *General coding decisions*
- 1253 Supplementary File 2B
- 1254 *Study-specific coding decisions*
- 1255
- 1256 Supplementary File 3A
- 1257 General moderators: for all predictors
- 1258 Supplementary File 3B
- 1259 Facial masculinity moderators
- 1260 Supplementary File 3C
- 1261 *Body masculinity moderators*
- 1262 Supplementary File 3D
- 1263 *2D:4D moderators*
- 1264 Supplementary File 3E
- 1265 Voice pitch moderators
- 1266 Supplementary File 3F
- 1267 Height moderators
- 1268 Supplementary File 3G
- 1269 *Testosterone levels moderators*
- 1270
- 1271 Supplementary File 4A
- 1272 Facial masculinity: moderation analyses. The intercept shows the 'simple effect' for the
- 1273 reference category (specified) and the moderator effect shows the change in effect size for
- 1274 that category relative to the reference category. Moderators are bolded if significant after

- 1275 controlling for multiple comparisons, as indicated by computation of q-values. The full list of
- 1276 *q-values can be found in Supplementary File 7.*
- 1277 Supplementary File 4B
- 1278 Body masculinity: moderation analyses. The intercept shows the 'simple effect' for the
- 1279 reference category (specified) and the moderator effect shows the change in effect size for
- 1280 *that category relative to the reference category. Moderators are bolded if significant after*
- 1281 *controlling for multiple comparisons, as indicated by computation of q-values. The full list of*
- 1282 *q-values can be found in Supplementary File 7.*
- 1283 Supplementary File 4C
- 1284 2D:4D: moderation analyses. The intercept shows the 'simple effect' for the reference
- 1285 *category (specified) and the moderator effect shows the change in effect size for that category*
- 1286 relative to the reference category. Moderators are bolded if significant after controlling for
- 1287 *multiple comparisons, as indicated by computation of q-values. The full list of q-values can be*
- 1288 found in Supplementary File 7.
- 1289 Supplementary File 4D
- 1290 Voice pitch: moderation analyses. The intercept shows the 'simple effect' for the reference
- 1291 *category (specified) and the moderator effect shows the change in effect size for that category*
- 1292 relative to the reference category. Moderators are bolded if significant after controlling for
- 1293 *multiple comparisons, as indicated by computation of q-values. The full list of q-values can be*
- 1294 found in Supplementary File 7.
- 1295 Supplementary File 4E
- 1296 Height: moderation analyses. The intercept shows the 'simple effect' for the reference
- 1297 *category (specified) and the moderator effect shows the change in effect size for that category*
- 1298 relative to the reference category. Moderators are bolded if significant after controlling for
- 1299 *multiple comparisons, as indicated by computation of q-values. The full list of q-values can be*
- 1300 *found in Supplementary File 7.*

1301 Supplementary File 4F

1302 *Testosterone levels: moderation analyses. The intercept shows the 'simple effect' for the*

- 1303 reference category (specified) and the moderator effect shows the change in effect size for
- 1304 *that category relative to the reference category. Moderators are bolded if significant after*
- 1305 controlling for multiple comparisons, as indicated by computation of q-values. The full list of
- 1306 *q-values can be found in Supplementary File 7.*
- 1307
- 1308 Supplementary File 5A
- 1309 Mating domain, reproductive domain, and offspring mortality domain predicted by global
- 1310 masculinity. Pearson's r (95% CI); p value for meta-analytic effect, q-value (correcting for
- 1311 *multiple comparisons); number of observations (k), samples (s), and unique participants (n);*
- 1312 *test for heterogeneity (Q), p value for heterogeneity. Statistically significant meta-analytic*
- 1313 associations are bolded if still significant after controlling for multiple comparisons.
- 1314 Supplementary File 5B
- 1315 *Global masculinity: moderation analyses. The intercept shows the 'simple effect' for the*
- 1316 *reference category (specified) and the moderator effect shows the change in effect size for*
- 1317 *that category relative to the reference category. Moderators are bolded if significant after*
- 1318 controlling for multiple comparisons, as indicated by computation of q-values. The full list of
- 1319 *q-values can be found in Supplementary File 7.*
- 1320
- 1321 Supplementary File 6A
- 1322 Funnel plots of effect sizes for mating measures (MAT). T = testosterone.
- 1323 Supplementary File 6B
- 1324 *Funnel plots of effect sizes for reproductive measures (REP). T* = *testosterone levels.*
- 1325
- 1326 Supplementary File 7

Output for q-value computation for all analyses

Authors	Year	Predictor	Outcome	Sample	Sample location	Low or high fert.	N
Alvergne et al.	2009	Т	REP	Rural villagers	Senegal	High	53
Apicella	2014	Body mase	MAT, REP, OM	Hadza	Tanzania	High	51
Apicella et al.	2007	Body masc, voice pitch, height	REP, OM	Hadza	Tanzania	High	44-52
Arnocky et al.	2018	Facial masc	MAT	Students	Canada	Low	135
Aronoff et al.	2017	Т	MAT	Students	US	Low	99
Atkinson	2012	Body masc	MAT	Students	US	Low	66
Atkinson et al.	2012	Body masc, voice pitch, height	REP	Himba (Ovahimba)	Namibia	High	36
Bogaert & Fisher	1995	Т	MAT	Students	Canada	Low	195-196
Booth et al.	1999	Т	MAT	Army veterans and non-veterans	US	Low	4393
Boothroyd et al.	2008	Facial masc	MAT	Students	UK	Low	18-19
Boothroyd et al.	2011	Facial masc	MAT	Students	UK	Low	36
Boothroyd et al.	2017	Facial mase	REP, OM	Agta	Philippines	High	65
		Facial masc	MAT, REP, OM	Maya	Belize	High	23-35
Charles & Alexander	2011	2D:4D, T	MAT	Students	US	Low	25-42
Chaudhary et al.	2015	Body masc, height	MAT, REP, OM	Mbendjele BaYaka	Democratic Republic of the Congo	High	55-73
Edelstein et al.	2011	Т	MAT	Students	US	Low	134
Falcon	2016	2D:4D	MAT	Students	US	Low	137
Farrelly et al.	2015	T	MAT	Students	UK	Low	75-78
Frederick	2010	Body masc, 2D:4D, height	MAT	Students	US	Low	61
Frederick & Haselton	2007	Body mase	MAT	Students	US	Low	56-121
Frederick & Jenkins	2015	Height	MAI	Online	Worldwide	Low	28/59-31418
Ganup et al.	2007	Body mase, 2D:4D	DED	Sudenis Former teenege delinguents	US	Low	101
Gettler et el	2008	T Dody mase	KEP MAT	Cobu Longitudinal Health and Nutrition Survey	Dhilipping	Lich	101
Gildnor	2019	I Rody mass 2D:4D height	DED	Shuar Health and Life History Project	Fander	High	200
Gómez-Valdés et al	2013	Facial mase	REP	Hallstatt skulls	Austria	High	170
Hartl et al	1982	Body mase height	MAT REP	Former teenage delinquents	US	High	180-185
Hill et al	2013	Facial mase, body mase.	MAT	Students	US	Low	63
	2010	voice pitch, height				2011	
Hoppler, Walther et al.	2018	Т	REP	Men's health 40+ study	Switzerland	Low	268
Hughes & Gallup Jr.	2003	Body mase	MAT	Students	US	Low	50-59
Hönekopp et al.	2006	2D:4D, height	MAT	Students and non-students	Germany	Low	79-99
Hönekopp et al.	2007	Facial mase, body mase, height. T	MAT	Students and non-students	Germany	Low	77
Kirchengast	2000	Height	REP, OM	!Kung San	Namibia	High	103
Kirchengast & Winkler	1995	Height	REP, OM	Urban and rural Kavango people	Namibia	High	59-78
Klimas et al.	2019	Т	MAT	Men's health 40+ study	Switzerland	Low	159
Klimek et al.	2014	2D:4D, height	REP	Mogielica Human Ecology Study Site	Poland	High	238
Kordsmeyer et al.	2018	Body masc, voice pitch, height, T	MAT	Students and non-students	Germany	Low	103-164
Kordsmeyer & Penke	2017	2D:4D, height	MAT	Students and non-students	Germany	Low	141
Krzyzanowska et al.	2015	Height	REP	National Child Development Study	UK	Low	6535
Kurzban & Weeden	2005	Height	MAT, REP	Speed daters	US	Low	1503-1501
Lassek & Gaulin	2009	Body masc, height	MAT	NHANES III	US	Low	4167-5159
Little et al.	1989	Height	REP, OM	Rural; growth stunted	Mexico	High	103
Loehr & O'Hara	2013	Facial mase	REP	WWII soldiers	Finland	High	795
Longman et al.	2018	T	MAT	Students	UK	Low	38
Luevano et al.	2018	Facial masc, height	MAT	Students	US	Low	35-66
Lukaszewski et al.	2014	Body masc	MAT	Students	US	Low	48-174
Maestripieri et al.	2014	T	MAT	Students	US	Low	41-61
Manning & Fink	2008	2D:4D	MAT, REP	Online	Worldwide	Low	26872-83681
Manning et al.	2003	2D:4D	REP	Community	England	Low	189
		2D:4D	REP	Sugali and Yanadi tribal groups	India Couth A fries	High	80
Marazalz at al	2019	2D:4D	REP	Zulus from townships near Durban	Judonosio	High	00
Maleture et al.	2018	2D.4D T	KEP MAT	I all Students	TIGOLIESIA	Low	47
Međedović & Bulut	2000	Height	MAT	Students	Serbia	Low	30
Mosing et al	2015	Height	MAT REP	Study of Twin Adults: Genes and Environment	Sweden	Low	2310-2549
Mueller & Mazur	1997	Facial mase	REP	West Point class of 1950	US	High	337
Nagelkerke et al	2006	Height	MAT	NHANES 99-00	US	Low	798-809
Nettle	2002	Height	REP	National Child Development Study	UK	Low	4474
Pawlowski et al.	2008	Height	REP	Rural	Poland	High	46
Pawlowski et al.	2000	Height	REP	Urban and rural	Poland	High	3201
Peters et al.	2008	Facial mase, body mase, T	MAT	Students	Australia	Low	100-113
Pollet et al.	2011	Т	MAT	National Social Life, Health, and Aging Project	US	Low	749
Polo et al.	2019	Facial mase, body mase, height	MAT	Students and non-students	Chile	Low	198-206
Price et al.	2013	Body masc, height	MAT	Mainly students	UK	Low	55
Prokop & Fedor	2011	Height	REP	Friends and family of students	Slovakia	Low	499
Prokop & Fedor	2013	Height	MAT	Students	Slovakia	Low	105-150
Puts et al.	2006	Voice pitch	MAT	Students	US	Low	103
Puts et al.	2015	Τ	MAT	Students	US	Low	59-61
Putz et al.	2004	2D:4D	MAT	Students	US	Low	207-219
Rahman et al.	2005	2D:4D, height	MAT	Students and non-students	UK	Low	78-150
Rhodes et al.	2005	Facial masc, body masc, height	MAT	Mainly students	Australia	Low	142-166
Rosenfield et al.	2020	Body mase, voice pitch, height	MAT, REP, OM	Isimané	Bolivia	High	55-62
Schwarz et al.	2011	2D:4D	MAI	Students	Germany	Low	52-89
Shown & Caller	1982	Rody mass 2D:4D	KEP MAT	Students		Low	28.29
Shoup & Gallup	2008	Body mase, 2D:4D	MAT	Students	US	Low	28-38
Simmons & Donay	2010	Body mass, 2D.4D	MAT	Students		Low	128
Smith et al	2011	Body mase	DED	Hadza	Tanzania	High	51
Sneade & Furnham	2017	Body mase	MAT	Students	IIK	Low	145
Sorokowski et al	2013	Height	REP OM	Yali	Indonesia	High	49-52
Steiner et al	2013	2D·4D T	REP	Students and non-students	US	Low	30
Stern et al.	2020	T	MAT	Students	UK	Low	61
Strong et al.	2014	Body mase	MAT	Students	US	Low	31
Strong & Luevano	2014	Body mase, 2D:4D, height	MAT	Students	US	Low	51-66
Subramanian et al.	2009	Height	OM	2005-2006 National Family Health Survey	India	Low	21120
Suire et al.	2018	Voice pitch	MAT	Mainly students	France	Low	57-58
Tao & Yin	2016	Height	REP	The Panel Study of Family Dynamics	Taiwan	Low	1409
van Anders et al.	2007	Т	MAT	Non-students	US	Low	31
Van Dongen & Sprengers	2012	Facial mase, body mase, 2D:4D	MAT	Not specified	Not specified	Low	52
Varella et al.	2014	Body masc, 2D:4D, height	MAT	Students	Brazil, Czech Republic	Low	69-80
von Rueden et al.	2011	Body masc, height	REP, OM	Tsimané	Bolivia	High	162-197
Voracek et al.	2010	2D:4D, height	REP	Firefighters	Austria	Low	134
Walther et al.	2016	Body masc	REP	Men's health 40+ study	Switzerland	Low	271
Walther et al.	2017a	Body masc	MAT	Men's health 40+ study	Switzerland	Low	226
Walther et al.	2017b	Height	REP	Men's health 40+ study	Switzerland	Low	271
Walther et al.	2017c	Height	MAT	Men's health 40+ study	Switzerland	Low	226
Waynforth	1998	2D:4D, height	MAT, REP, OM	Villagers	Belize	High	35-56
Weeden & Sabini	2007	Body mase, 2D:4D, height	MAT	Students	US	Low	188-212
winkler & Kirchengast	1994	Height	KEP, OM	!Kung San	Namibia	High	51-114

Note. Fert = fertility, masc = masculinity, MAT = mating, NHANES = National Health and Nutrition Examination Survey, OM = offspring mortality, REP = reproduction, T = testosterone.

Overall analyses (all traits & all samples combined)

- Global masculinity (all masculine traits) predicting:
 - Total fitness
 - Mating domain
 - Reproductive domain
 - Offspring viability domain

Main analyses (separating masculine traits; all samples combined)

- Each masculine trait predicting:
 - Mating domain
 - Reproductive domain

Subgroup analyses (separating sample type & outcome measure type)

- Low fertility samples: Each masculine trait predicting:
 - Mating domain
 - Mating attitudes
 - Mating behaviors
 - Reproductive domain
 - Fertility
 - Reproductive success
- High fertility samples: Each masculine trait predicting:
 - Mating domain
 - -
 - Mating attitudes
 - Mating behaviors
 - Reproductive domain
 - Fertility
 - Reproductive success

Moderation analyses (full details in Supplementary Files 3)

- Domain type (mating vs reproduction)
 - Mating measure type (attitudes vs behaviors)
 - Reproductive measure type (fertility vs reproductive success)
- Sample type (low vs high fertility)
 - Low fertility sample type (student vs non-student sample)
 - o High fertility sample type (traditional vs industrialized sample)
- Ethnicity
- Marriage system
- Publication type
- Peer-review status
- Sexual orientation
- Transformation of variables
- Conversion of effect size
- Age control
- Other control variables

Author	(s)	and	year
	(-/		J

Gallup et al., 2007 Gallup et al., 2007 Gallup et al., 2007 Gallup et al., 2007 Varella et al., 2014 Shoup & Gallup, 2008 Shoup & Gallup, 2008 Van Dongen & Sprengers, 2012 Van Dongen & Sprengers, 2012 Sneade & Furnham, 2016 Sneade & Furnham, 2016 Sneade & Furnham, 2016 Frederick, 2010 Strong et al., 2014 Strong et al., 2014 Strong et al., 2014 Strong et al., 2014 Strong & Luevano, 2014 Kordsmeyer et al., 2018 Kordsmeyer et al., 2018 Kordsmeyer et al., 2018 Chaudhary et al., 2015 Lukaszewski et al., 2014 Hartl et al., 1982 Atkinson, 2012 Atkinson, 2012 Atkinson, 2012 Gallup et al., 2007 Gallup et al., 2007 Sim & Chun, 2016 Sim & Chun, 2016 Shoup & Gallup, 2008 Shoup & Gallup, 2008 Hughes & Gallup Jr., 2003 Price et al., 2013 Price et al., 2013 Price et al., 2013 Price et al., 2013 Kordsmeyer et al., 2018 Kordsmeyer et al., 2018 Kordsmeyer et al., 2018 Varella et al., 2014 Price et al., 2013 Price et al., 2013 Price et al., 2013 Price et al., 2013 Weeden & Sabini, 2007 Weeden & Sabini, 2007 Apicella, 2014 Frederick & Haselton, 2007 Lassek & Gaulin, 2009 Hill et al., 2013 Lukaszewski et al., 2014 Rhodes et al., 2005 Rhodes et al., 2005



71	0.25 [0.01,	0.49]
75	0.34 [0.11.	0.571
71	0.24[0.00]	0 481
75		0.51
75	0.20[0.05,	0.51
80	0.36 [0.13,	0.58]
69	0.31 [0.06,	0.55]
80	0.00 [-0.22,	0.22]
80	0.00 [-0.22,	0.22]
80	0.00[-0.22]	0.221
38	0.33[-0.00	0.661
30 04	0.00 [0.00, 0.00]	0.00
31	0.02[-0.35]	0.39]
52	0.09[-0.19,	0.37]
52	-0.04 [-0.32,	0.24]
145	0.22 [0.06,	0.39]
145	0.16 [-0.00,	0.33]
145	0.27 [0.10,	0.43]
61	0.15 [-0.11,	0.41]
31	0 00 [-0 37	0 371
21	0.00[-0.07]	0.601
31	0.23[-0.14]	0.00]
31	-0.01[-0.38,	0.30
31	0.04 [-0.33,	0.41]
51	-0.14 [-0.42,	0.14]
66	0.05 [-0.20,	0.30]
66	-0.02 [-0.27,	0.23]
66	0.15 [-0.10,	0.40]
66	-0.09[-0.34]	0.161
16/	0.15[-0.00]	0.311
104	0.10[0.00]	0.01
164	0.13[-0.02,	0.29]
104	0.09 [-0.10,	0.29]
70	0.05 [-0.19,	0.29]
174	0.10 [-0.05,	0.25]
170	0.16 [0.01,	0.32]
165	0.25 [0.09.	0.401
7/	021[-002]	0 4 4 1
74	0.21[0.02, 0.14[-0.00]	0.77]
74	0.14[-0.09]	0.37]
73	0.12[-0.12,	0.35]
48	0.28 [-0.01,	0.57]
48	-0.02 [-0.31,	0.28]
108	0.20[0.01,	0.40]
184	0.23 [0.08,	0.37]
66	0.33[0.08]	0.581
66	0.00[0.00]	0.00]
00	0.24[0.01,	0.40
66		0.50]
71	0.06 [-0.17,	0.30]
75	0.14 [-0.09,	0.37]
90	0.10 [-0.11,	0.31]
90	-0.04 [-0.25,	0.17]
38	0.23 [-0.10,	0.56]
31	0.17 [-0.20,	0.55]
50	0 29 [0 00	0 571
50	0.50[0.00,	0.761
59		0.70
59	0.27 [0.00, 0.20]	0.001
59	0.36[0.10,	0.63]
55	0.06 [-0.22,	0.33]
55	0.11 [-0.16,	0.38]
55	0.01 [-0.26,	0.28]
55	0.06 [-0.21,	0.33]
164	-0.07 [-0.22,	0.08]
164	-0.05 [-0.20,	0.10]
104	-0.07 [-0.27,	0.12]
80	0.00[-0.22]	0.221
80 80	0.00[-0.22]	0.221
00 00	0.00[-0.22]	0.221
00	0.00 [0.22, 0.02]	0.22]
80		0.22]
80	0.00 [-0.22,	0.22]
80	0.00 [-0.22,	0.22]
55	0.09 [-0.18,	0.36]
55	-0.02 [-0.29,	0.25]
55	-0.05 [-0.32,	0.22]
55	-0.02 [-0.29,	0.25]
197	0.02 [-0.12,	0.16]
197	0.04 [-0.10,	0.18]
51	0.25 [-0.03.	0.531
gg	0.09[-0.11]	0 291
101		0.461
121	0.20[0.10,	0.40]
112		0.39]
56	0.26 [-0.01,	0.52]
56	0.33 [0.06,	0.60]
56	0.28 [0.01,	0.55]
4774	0.11 [0.08,	0.14]
4737	0.05 [0.02,	0.08]
4167	-0.05 [-0.08, -	-0.02]
4774	0.05 [0.02,	0.08]
4737	0.13 [0.10.	0.161
4167	0.05[0.01]	0.081
63	0.68[0.42]	0 931
17/	0.22 [0.72, 0.22 [0.72	0 271
170	0.25 0.07	0.401
170	$0.25 \downarrow 0.09,$	0.40]
00	$0.23 \downarrow 0.10,$	0.41]
/4 	0.20[0.03,	0.49
<i>/</i> 4	0.13 [-0.10,	0.36]
73	0.21 [-0.02,	0.45]
48	0.14 [-0.16,	0.43]
48	-0.10 [-0.40,	0.19]
108		0 371
100	0.18[-0.01,	0.07]
142	0.18 [-0.01, 0.12 [-0.05,	0.29]

Fisher's Z

Ν



Author(s) and year

Hill et al., 2013 Puts et al., 2006 Kordsmeyer et al., 2018 Kordsmeyer et al., 2018 Kordsmeyer et al., 2018 Suire et al., 2018 Suire et al., 2018 Rosenfield et al., 2020

RE Model

Ν



Fisher's Z

0.07 [-0.18, 0.32] 0.15 [-0.04, 0.35] 0.16 [0.01, 0.32] 0.18 [0.03, 0.34] 0.09 [-0.11, 0.29] -0.02[-0.28, 0.25]0.12 [-0.14, 0.39] 0.17 [-0.10, 0.45]

0.13 [0.06, 0.20]

Author	(s)	and	vear
Aution	(9)	and	ycar

Charles & Alexander, 2011 Charles & Alexander, 2011 Charles & Alexander, 2011 Bogaert & Fisher, 1995 Bogaert & Fisher, 1995 Edelstein et al., 2011 Edelstein et al., 2011 Edelstein et al., 2011 Farrelly et al., 2015 Maestripieri et al., 2014 Maestripieri et al., 2014 Maestripieri et al., 2014 Maestripieri et al., 2014 Booth et al., 1999 Aronoff et al., 2017 Aronoff et al., 2017 Aronoff et al., 2017 Aronoff et al., 2017 van Anders et al., 2007 McIntyre et al., 2006 McIntyre et al., 2006 Pollet et al., 2011 Puts et al., 2015 Peters et al., 2008 Simmons & Roney, 2011 Simmons & Roney, 2011 Kordsmeyer et al., 2018 Kordsmeyer et al., 2018 Kordsmeyer et al., 2018 Hönekopp et al., 2007 Hönekopp et al., 2007 Hönekopp et al., 2007 Gettler et al., 2019 Gettler et al., 2019 Gettler et al., 2019 Gettler et al., 2019 Gettler et al., 2019



Ν	Fisher's Z
42	-0.23 [-0.55, 0.08]
42	-0.21 [-0.53, 0.10]
12	-0.24[-0.56, 0.07]
т <u>с</u> 105	0.20[0.06, 0.07]
195	0.20[0.00, 0.04] 0 1/ [_0 00 0 28]
196	0.14[-0.00, 0.20]
134	0.17 [0.00, 0.34]
134	0.06 [-0.11, 0.23]
134	0.20[0.03, 0.37]
78	0.08 [-0.15, 0.31]
78	0.05 [-0.18, 0.28]
78	0.18 [-0.04, 0.41]
75	0.15 [-0.08, 0.38]
75	0.03 [-0.20, 0.26]
61	0.07 [-0.19, 0.33]
61	0.03 [-0.23, 0.29]
45	0.06 [-0.24, 0.36]
41	0.02 [-0.30, 0.34]
4393	0.07[0.04.0.10]
99	0.09[-0.11.0.29]
gg	0.09[-0.11_0.29]
00 00	0.08[-0.12, 0.28]
99	0.00[0.12, 0.20] 0.02[-0.18, 0.22]
99 04	0.02 [0.10, 0.22] 0.02 [0.10, 0.22]
3 I 0 4	0.00[-0.29, 0.40]
81	0.04 [-0.10, 0.20]
68	0.11[-0.13, 0.35]
749	0.12[0.05, 0.19]
59	0.34 [0.08, 0.60]
59	0.32 [0.06, 0.58]
59	-0.02 [-0.28, 0.24]
59	-0.06 [-0.32, 0.20]
59	0.30 [0.04, 0.56]
61	0.21 [-0.04, 0.47]
61	0.16 [-0.10, 0.42]
61	-0.01 [-0.27, 0.25]
61	-0.16 [-0.42, 0.10]
61	0.16 [-0.10, 0.42]
107	0.17 [-0.02, 0.37]
107	0.19 [-0.00, 0.38]
107	0.08 [-0.11, 0.27]
100	0.03 [-0.16, 0.23]
100	-0.00[-0.19, 0.29]
122	0.00[-0.10, 0.10] 0.08[-0.09, 0.25]
100	
130	-0.01[-0.10, 0.10] 0.12[-0.02, 0.20]
163	
163	
104	0.09[-0.10, 0.29]
17	0.00[-0.23, 0.23]
77	0.30 [0.08, 0.53]
77	0.11 [-0.12, 0.34]
288	0.09 [-0.03, 0.21]
288	0.06 [-0.06, 0.18]
288	0.03 [-0.09, 0.15]
288	0.08 [-0.04, 0.20]
288	0.07 [-0.05, 0.19]
000	



Author(s) and year	Ν	Fisher's Z
Varella et al., 2014	80	0.00 [-0.22, 0.22]
Varella et al., 2014	80	0.00 [-0.22, 0.22]
Varella et al., 2014	80	0.00 [-0.22, 0.22]
Varella et al., 2014	80	0.00 [-0.22, 0.22]
Varella et al., 2014	80	0.00 [-0.22, 0.22]
Varella et al., 2014	80	0.00 [-0.22, 0.22]
Frederick & Jenkins, 2015	31418	0.12 [0.11, 0.13]
Frederick & Jenkins, 2015	28759	0.02 [0.01, 0.04]
Hill et al., 2013	63	-0.18 [-0.44, 0.07]
Rhodes et al., 2005	143	0.20 [0.04, 0.37]
Rhodes et al., 2005	165	0.18 [0.03, 0.34]
Rhodes et al., 2005	– – 160	0.26 [0.10, 0.42]
Rhodes et al., 2005	164	-0.02 [-0.18, 0.13]
Rahman et al., 2005	150	0.17 [0.01, 0.33]
Frederick, 2010	- 61	0.27 [0.01, 0.53]
Price et al., 2013	55	0.09 [-0.18, 0.36]
Price et al., 2013	55	0.18 [-0.09, 0.45]
Price et al., 2013	55	0.11 [-0.16, 0.38]
Price et al., 2013	55	0.17 [-0.10, 0.45]
Nagelkerke et al., 2006	809	-0.04 [-0.11, 0.03]
Nagelkerke et al., 2006	798	0.07 [0.00, 0.14]
Nagelkerke et al., 2006	806	0.02 [-0.05, 0.08]
Kurzban & Weeden, 2005	1504	0.06[0.01.0.11]
Wavnforth, 1998	<u> </u>	0.23 [-0.04, 0.50]
Kordsmever & Penke. 2017	141	-0.10 [-0.27, 0.07]
Kordsmever & Penke, 2017	141	-0.13 [-0.30, 0.04]
Kordsmever & Penke, 2017	141	-0.07 [-0.24, 0.10]
Kordsmever & Penke, 2017	141	-0.08 [-0.25, 0.09]
Kordsmever et al., 2018		0.12 [-0.03, 0.28]
Kordsmever et al., 2018	164	0.13 [-0.02, 0.29]
Kordsmever et al., 2018	104	0.07 [-0.12, 0.27]
Lassek & Gaulin, 2009	5159	0.05[0.02,0.07]
Lassek & Gaulin, 2009	5093	0 02 [-0 01 0 05]
Lassek & Gaulin, 2009	5092	0.09[0.06, 0.11]
Weeden & Sabini 2007		0.12[-0.01, 0.26]
Weeden & Sabini, 2007	212	0.12[-0.02, 0.25]
Prokon & Fedor 2013		0.11[-0.02, 0.20] 0.10[-0.06, 0.26]
Prokon & Fedor 2013	105	
Prokop & Fedor, 2013	105	0.20 [0.00, 0.00]
$\frac{100000 \text{ a Fedor, 2013}}{2013}$	ー 149 150	
$\frac{1}{10000000000000000000000000000000000$		
Luevano et al., 2018/Strong & Luevano, 2014 $\begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$	ר כר הכר	
Luovano et al. 2018/Strong & Luovano, 2014	00 	-0.04 [-0.29, 0.21] -0.01 [-0.26, 0.24]
Luevano et al., 2010/Strong & Luevano, 2014	Ч ОО СС	
Luevano et al., 2018/Strong & Luevano, 2014		0.04 [-0.21, 0.29] 0.01 [-0.24, 0.26]
Luevano et al., 2010/Shong & Luevano, 2014 $\begin{bmatrix} -1 \\ -1 \end{bmatrix}$	66	
Hönekopp et al., 2000 $ $	93	0.04 [-0.10, 0.25]
$\square Onekopp et al., 2007$	//	
$\square One KOpp et al., 2007$	77	-0.09[-0.32, 0.14]
Unaudhary et al., 2015 Maairen at al. 2015	73	-0.08[-0.32, 0.15]
Nosing et al., 2015	2321	-0.00[-0.04, 0.04]
Nosing et al., 2015 Masimum et al., 2015	2549	
	0100	<u> </u>





Author(s) and year

Atkinson et al., 2012 Chaudhary et al., 2015 Chaudhary et al., 2015 Hartl et al., 1982 Gildner, 2018 Apicella, 2014 Smith et al., 2017 Genovese, 2008 Gildner, 2018 Gildner, 2018 von Rueden et al., 2011 Rosenfield et al., 2020 Rosenfield et al., 2020 Walther et al., 2016

RE Model

-0.5





0.14 [0.08, 0.21]