SY-04 Replication map: what holds and what fell

The replication crisis in the psychological and biological sciences has precipitated a fundamental epistemological shift in how human behavior is understood. Findings that once served as the bedrock of textbooks have been subjected to unprecedented scrutiny through large-scale, pre-registered replications, mega-analyses, and adversarial collaborations. Nowhere is this scrutiny more consequential—or more culturally volatile—than in the study of sex differences and hormone-behavior interactions. Exploring how men and women differ, and how they are alike, is a domain frequently complicated by ideological commitments, measurement artifacts, and publication bias.

To systematically evaluate the evidence, it is necessary to construct a taxonomy of replication that distinguishes between what is known about statistical population averages and what can be predicted about any specific individual. Because human psychological and behavioral traits are continuously distributed, male and female distributions overlap substantially. Consequently, a statistically robust average difference between groups—a shift in the mean—does not preclude vast intra-sexual variation, nor does it reliably predict individual outcomes. The average man and the average woman are statistical aggregates; the individual is a unique configuration of traits.

To provide a rigorous foundation for understanding human group dynamics, this report maps the replication status of major claims regarding sex differences. The analysis separates findings into three distinct but constantly interacting layers: physiology (measurable biological parameters), behaviour (statistical distributions of traits and actions), and culture (variable social structures and norms).

The Physiological Layer: Measurable Biology

The physiological layer encompasses anatomical, neurobiological, and endocrinological differences between the sexes. While this layer is often assumed to be the most objective, it remains a site of fierce methodological debate regarding how variables are measured, scaled, and interpreted.

Physical Strength and the Facultative Calibration of Personality

Sex differences in physical strength represent some of the largest, most visible, and most robust physiological effect sizes documented in the human species. A comprehensive meta-analysis of physical ability tests used in occupational contexts revealed that males score substantially higher on tests of muscular strength and cardiovascular endurance[1]. Specifically, upper-body strength yields a massive effect size, routinely exceeding a Cohen's d of 2.0, meaning the male mean is more than two standard deviations above the female mean[2]. Females produce, on average, 60% to 75% of male absolute upper-extremity strength[1]. While these absolute differences diminish when controlling for lean body mass and height, they do not disappear entirely[3].

The replication status of absolute physical strength differences is undeniably robust. Moreover, this physiological divergence has been theorized to mediate other behavioral traits. Manson et al. (2022, *Social Psychological and Personality Science*) conducted a study across five samples ($N = 1,399$) testing the hypothesis that physical strength calibrates fear and anxiety. They found that physical strength statistically mediated the sex difference in the HEXACO Emotionality factor, particularly its Fearfulness and Anxiety facets[4]. This suggests that evolutionary mechanisms may facultatively calibrate a person's level of threat avoidance in response to their internal estimate of relative physical form[4].

Neuroanatomy: Dimorphism versus the Brain Mosaic

The question of whether male and female brains differ structurally has undergone significant empirical re-evaluation, heavily influenced by sample size and statistical methodology. For years, functional magnetic resonance imaging (fMRI) studies relying on small sample sizes produced highly localized, often exaggerated claims of sex differences in brain activation. Most of these small-sample fMRI claims have failed to replicate, becoming prime casualties of the replication crisis[5].

However, macro-structural differences hold up under scrutiny. Ruigrok et al. (2014, *Neuroscience & Biobehavioral Reviews*) conducted a meta-analysis showing that males possess, on average, a 10% to 13% larger total brain volume[7]. Adjusting for this total volume difference, specific regional differences emerge: males show greater average grey matter volume in the amygdala, hippocampus, and putamen, while females exhibit higher cortical thickness and greater white matter tract complexity[5]. These findings have been reliably replicated in mega-samples, such as the UK Biobank imaging data ($N > 5,000$), yielding a median absolute standardized effect size of $d = 0.17$ for structural phenotypes[6].

Despite these robust structural findings, the interpretation remains highly controversial. Joel et al. (2015, *Proceedings of the National Academy of Sciences* [*PNAS*]) analyzed MRIs from over 1,400 brains and proposed the "human brain mosaic" hypothesis. They argued that because the presence of male-typical and female-typical structures varies independently within the same brain, individual brains cannot be strictly categorized into two distinct, dimorphic classes[10]. In response, Del Giudice et al. (2016, *PNAS*) critiqued this methodology, demonstrating that when multivariate statistical techniques are applied to Joel et al.'s exact dataset, individuals can be correctly classified by sex with high accuracy (up to 77%), refuting the claim that male and female brain structures are indistinguishable[10]. The physiological differences are robust, but the theoretical classification paradigm—whether the brain is a mosaic of overlapping traits or a multivariate dimorphism—remains deeply contested.

Hormones and Behavior: The Fatherhood Testosterone Drop

A cornerstone of evolutionary endocrinology is the trade-off between mating effort (facilitated by high testosterone) and parenting effort. Gettler et al. (2011, *PNAS*) provided definitive longitudinal evidence demonstrating this trade-off in human males. Tracking a cohort of 394 men over five years in the Philippines, the researchers found that men with high baseline

testosterone were more likely to become partnered fathers; however, upon becoming fathers, they experienced a precipitous drop in testosterone compared to men who remained single and childless[14].

This decline in testosterone is interpreted as an adaptive, physiological mechanism designed to facilitate caregiving and reduce aggressive mating competition[17]. Subsequent replications and broader systemic reviews have confirmed that fatherhood, particularly engaged and proximate parenting, is associated with reduced testosterone across multiple cultures. This physiological shift is considered a robust, highly replicated finding in human behavioral endocrinology[17].

The Retraction of Ovulatory Mate-Preference Shifts

For over a decade, evolutionary psychology literature was dominated by the claim that women experience distinct shifts in mate preferences across their ovulatory cycle—specifically, that during peak fertility, women strongly prefer highly masculine, symmetrical, and dominant men. However, recent large-scale, pre-registered mega-analyses and longitudinal studies utilizing precise hormonal tracking (rather than counting days on a calendar) have completely failed to replicate these effects. The claim of ovulatory mate-preference shifts is now widely considered retracted or failed, serving as a cautionary tale of p-hacking and publication bias in early 2000s psychological literature[21].

The Behavioural Layer: Statistical Distributions

The behavioural layer encompasses cognitive traits, personality, vocational interests, and empathy. Because these traits are measured via psychometrics and self-reports, they are inherently statistical, characterized by overlapping Gaussian distributions. The leading debate in this layer concerns whether sex differences are trivial or massive—a discrepancy driven almost entirely by the choice between univariate and multivariate statistical methodologies.

Personality: The Univariate versus Multivariate Debate

In 2005, Hyde published the highly influential "gender similarities hypothesis" in the *American Psychologist*, synthesizing 46 meta-analyses (encompassing millions of data points) to argue that males and females are overwhelmingly similar on most psychological variables. Hyde noted that 78% of reported univariate effect sizes (Cohen's d) were small ($d < 0.35$) or near zero[22]. For example, univariate sex differences in the Big Five personality traits are modest: neuroticism ($d = -0.49$) and agreeableness ($d = -0.22$) are higher in women, while openness shows negligible differences[25]. Hyde cautioned that overinflated claims of gender differences carry substantial societal costs, fostering stereotyped expectations in the workplace and in relationships[23].

However, this paradigm was aggressively challenged by Del Giudice, Booth, and Irwing (2012, *PLoS ONE*). The leading critics argued that evaluating personality one trait at a time fundamentally obscures the biological and psychological reality of how traits cluster together in actual human beings. Utilizing a large, representative U.S. sample ($N = 10,261$) and the 16PF Questionnaire, they applied a multivariate effect size metric known as the Mahalanobis

distance (D). While univariate differences on single traits were indeed small (supporting Hyde's data), the aggregate multivariate difference between male and female personality profiles was extremely large ($D =$). This multivariate distance indicates an overlap of only 10% between the male and female multidimensional distributions[27].

This controversy does not represent a failure of replication; both Hyde and Del Giudice used robust data. Instead, it highlights a fundamental mathematical divergence. If one asks whether the average man and woman differ on any *single* isolated trait, the answer is "barely." If one asks whether the *overall profile* of an individual differs across a multi-trait space, the answer is "substantially." The raw data holds, but the choice of metric remains deeply ideological and context-dependent[25]. This perfectly illustrates the necessity of distinguishing between group averages and individual prediction: a single univariate average tells us almost nothing about an individual, but a multivariate profile can predict sex with high accuracy.

Vocational Interests: Men, Women, Things, and People

One of the most robust and widely replicated findings in differential psychology relates to vocational interests. Su, Rounds, and Armstrong (2009, *Psychological Bulletin*) conducted an exhaustive meta-analysis spanning 47 interest inventories and $N = 503,188$ respondents[32]. They found an exceptionally large sex difference on the Things–People dimension, with men demonstrating a strong preference for working with things and women demonstrating a strong preference for working with people, yielding a massive effect size of $d =$ [32].

Further granular analysis revealed that men consistently score higher on Realistic ($d = 0.84$) and Investigative ($d = 0.26$) interests, as well as specific STEM measures like engineering ($d = 1.11$). Conversely, women score higher on Social ($d = -0.68$) and Artistic ($d = -0.35$) dimensions[32]. These preferences demonstrate high cross-cultural stability and are considered a primary psychological mechanism underlying gender disparities in specific occupational fields, persisting even in societies with high gender equality. This finding is universally classified as robust[35].

Evolution and Mate Preferences

Evolutionary psychology predicts that due to asymmetrical obligate parental investment, women will prioritize cues of resource acquisition and willingness to invest, while men will prioritize cues of fertility (observable via youth and physical attractiveness). Walter et al. (2020, *Psychological Science*) conducted a massive 45-country replication ($N = 14,399$) of classic cross-cultural mate preference studies to test whether these evolutionary claims held up against modern biosocial role theories[37].

The replication successfully confirmed the original findings: universally, men prefer younger, physically attractive mates more than women do, and women prefer older mates with good

financial prospects more than men do[37]. However, the study also provided critical evidence for cultural modulation: as national gender equality increases, the age gap between preferred partners narrows, indicating that while foundational preferences are biologically tethered, their expression is environmentally calibrated[37]. The baseline sex differences in mate preferences are robust.

Empathy, Theory of Mind, and the Role of Motivation

The "Empathizing-Systemizing" theory posits that females, on average, have a stronger biological drive to empathize, while males have a stronger drive to systemize. Greenberg et al. (2018, *PNAS*) tested this using the "Reading the Mind in the Eyes" Test (RMET) on a colossal sample of over 500,000 individuals, confirming a persistent female advantage across lifespans and cultures[42].

However, the nature and magnitude of this difference are fiercely debated when shifting from capability to motivation. Klein and Hodges (2001, *Personality and Social Psychology Bulletin*) demonstrated that when male subjects were provided with a monetary incentive to perform well on empathetic accuracy tasks, the sex difference completely disappeared[45]. This suggests that what appears as a deficit in male cognitive empathy in standard psychological testing may actually be a deficit in *motivation* to deploy empathetic resources under artificial laboratory conditions[45]. Consequently, the claim of innate female superiority in empathy is considered mixed, heavily dependent on context, measurement, and incentive structures.

Sex Drive and Sociosexuality: Physiology versus Self-Report

Early foundational literature, such as Baumeister, Catanese, and Vohs (2001), concluded that men have a biologically stronger, more frequent, and more persistent sex drive than women, with a medium-to-large effect size[47]. Similarly, Schmitt (2005, *Behavioral and Brain Sciences*) found robust cross-cultural sex differences in sociosexuality (the willingness to engage in uncommitted, casual sex) across 48 nations, with men scoring significantly higher[50].

However, modern replications and methodological critiques have introduced severe caveats, highlighting the vulnerability of self-report metrics to societal scripts. Touraille and Ågmo (2024) demonstrated that when sexual motivation is measured via automatic physiological responses rather than self-report surveys, sex differences entirely disappear[47]. Furthermore, utilizing a "bogus pipeline" methodology—where participants believe they are attached to an infallible lie detector—the gender gap in self-reported sexual behaviors, such as masturbation frequency and casual sex, shrinks to statistical insignificance[47]. This suggests that societal gender norms and response bias heavily distort self-report sexual data. Therefore, innate differences in baseline sex drive and sociosexuality are graded as mixed to heavily caveated.

The Rise and Fall of Oxytocin

Early neurobiological studies labeled oxytocin the "cuddle hormone," claiming that intranasal administration broadly increased human trust, empathy, and general prosociality. However, large meta-analyses (e.g., Nave et al., 2015) subsequently failed to find any robust main effect of oxytocin on generalized trust[52].

Instead, the role of oxytocin was heavily revised by researchers like De Dreu et al. (2010, *Science*), who demonstrated that oxytocin actually promotes "parochial altruism." It increases

cooperation, cohesion, and empathy *only* for in-group members, while simultaneously increasing out-group derogation and ethnocentrism. Further studies by Zhang et al. (2019, *eLife*, $N = 480$) showed that oxytocin promotes coordinated out-group attack during intergroup conflicts[54]. The narrative of oxytocin as a universal prosocial hormone is retracted and failed, while its role as an in-group/out-group modulator is robust.

The Greater Male Variability Hypothesis

The Greater Male Variability Hypothesis (GMVH) suggests that males exhibit greater variance in phenotypic traits, being statistically overrepresented at both the extreme high and extreme low ends of a distribution, even when the mean is identical to females[57]. First posited by Charles Darwin and later expanded by Havelock Ellis, the hypothesis has found robust modern empirical support. Arden, Plomin, and Johnson (2010) established this in cognitive traits and brain structures, noting significantly greater variance in male cerebral white matter, cortex, and hippocampus volumes across development[57].

Extensive research also confirms greater male variability in physical strength, creative problem-solving, and general intelligence (often measured via Variance Ratios $>$)[57]. Despite historical sociological resistance to the concept, the evidence for GMVH across numerous physiological and psychometric domains is considered robust, potentially driven by evolutionary mechanisms related to X-linked genetic expression and sexual selection dynamics[58].

The Cultural Layer: Variable Structures and Paradoxes

The cultural layer involves how environments, economic policies, and societal structures interact with human psychology. This layer is highly susceptible to the replication crisis because societal interventions and correlational studies often fail to produce uniform results when applied across different cultural clusters.

The Gender-Equality Paradox

One of the most intensely debated phenomena in contemporary sociology and psychology is the "Gender-Equality Paradox" (GEP). Stoet and Geary (2018, *Psychological Science*; $N = 472,242$) analyzed global PISA data and discovered a highly counter-intuitive trend: as countries become more economically prosperous and gender-egalitarian (e.g., Nordic countries), the gender gap in STEM degree attainment actually *widens*[63]. Stoet and Geary explained this via expectancy-value theory and the resource hypothesis: in wealthy nations with a strong social safety net, students are free to pursue their intrinsic interests (which, per the Things-People divergence, differ by sex). In less developed nations, both sexes pursue STEM out of economic necessity, suppressing innate preferences[63]. Falk and Hermle (2018, *Science*; $N = 80,000$) documented a similar paradox, showing that gender gaps in altruism, trust, and risk-taking are larger in more egalitarian nations[67].

However, this biological-unmasking interpretation is actively contested by sociologists and

economists. Breda et al. (2020, *PNAS*; $N = 300,000$) demonstrated that the paradox can be entirely mediated by cultural stereotypes. They found that essentialist stereotypes (e.g., the belief that "math is for boys") are actually *stronger* in more developed and egalitarian countries. When statistically controlling for the prevalence of these specific stereotypes, the relationship between national wealth and the STEM gap becomes non-significant[69]. Thus, the paradox itself—that behavioral and occupational gaps widen in wealthy countries—is robust, but the proposed causal mechanism (unleashing innate biological preferences vs. the emergence of new horizontal gender norms) remains unknown and hotly contested[72].

Single-Sex Education

For decades, educational proponents argued that separating boys and girls in primary and secondary education improved academic outcomes by removing gendered distractions, mitigating stereotype threat, and allowing teachers to cater to "sex-specific" learning styles. Pahlke, Hyde, and Allison (2014, *Psychological Bulletin*) conducted a definitive meta-analysis of 184 studies representing 1.6 million students across 21 nations[74].

They found that while uncontrolled studies showed modest advantages for single-sex schooling, properly controlled studies (adjusting for socioeconomic status and prior achievement) yielded trivial differences ($g = 0.10$ for girls, $g = 0.06$ for boys in mathematics) and in some cases favored coeducational schooling[75]. The claim that single-sex education broadly improves academic performance is considered a failed hypothesis[77].

The Retraction of Power Posing

Carney, Cuddy, and Yap (2010, *Psychological Science*) famously claimed that holding expansive "power poses" for two minutes increased testosterone, decreased cortisol, and increased financial risk-taking. This became a cultural phenomenon, heavily promoted as a life-hack to empower women in the workplace. However, Ranehill et al. (2015, *Psychological Science*) conducted a rigorously controlled, pre-registered replication with a much larger sample size ($N = 200$) and found absolutely zero effect on testosterone, cortisol, or risk-taking, replicating only the subjective, self-reported feeling of power[79]. Subsequent mega-analyses confirmed these null behavioral and hormonal results, leading the original lead author (Carney) to publicly disavow the findings. The physiological and behavioral claims of power posing are retracted and failed[83].

Table 1: Graded Replication Map of Key Findings

Finding / Claim	Replication Status	Effect Size / Metric	Reason for Grade / Context
Physical Strength	**Robust**	Very Large ($d >$ upper	Consistently replicated across global cohorts;

		body)	fundamental anatomical dimorphism. Absolute gap narrows but persists when controlling for mass.[1]
Physical Aggression	**Robust**	Moderate to Large	Archer (2004) and subsequent replications confirm consistent male overrepresentation in direct physical violence across cultures.
Vocational Interests (Things vs. People)	**Robust**	Large ($d =$)	Replicated in mega-samples ($N >$). High cross-cultural stability; highly predictive of occupational sorting.[32]
Fatherhood & Testosterone Drop	**Robust**	Medium to Large	Longitudinal data (Gettler et al., 2011, $N =$) and cross-cultural replications confirm active caregiving lowers T in human males.[14]
Mate Preferences (Age/Resources)	**Robust**	Medium to Large	Walter et al. (2020) 45-country replication ($N =$) confirms

			evolutionary hypotheses, though equality slightly modulates the age gap size.[37]
Personality Differences (Multivariate)	**Robust**	Very Large ($D =$)	While univariate traits yield small gaps (Hyde 2005), multivariate clustering (Del Giudice 2012) robustly separates populations.[23]
Greater Male Variability (GMVH)	**Robust**	Variance Ratio $>$	Replicated in cognitive traits, intelligence, physical strength, and brain morphology (Arden et al. 2010).[57]
Oxytocin as "Prosocial/Cuddle " Hormone	**Mixed**	Small to Medium	Failed to replicate as universally prosocial. Robustly replicated as an in-group/out-group modulator (Parochial Altruism).[52]
Empathy / ToM Differences	**Mixed**	Small to Medium	RMET shows female advantage ($N >$), but monetary incentives entirely erase the gap. The gap appears largely motivational.[43]

Tend-and-Befriend	**Mixed**	Small	Female stress responses show more affiliative behavior than males, but hormonal mediators (oxytocin) have proven highly context-dependent.
Brain Mosaicism vs. Dimorphism	**Mixed / Unknown**	Varies by analytical model	Structural differences are robust (Ruigrok 2014); whether they constitute a true dimorphism or an overlapping mosaic is fiercely contested statistically.[5]
Gender-Equality Paradox Mechanism	**Unknown**	Medium	Paradox is real (occupational/preference gaps widen in rich nations), but the cause is contested: Biological trait unmasking vs. horizontal stereotypes.[63]
Ovulatory Mate-Preference Shifts	**Failed / Retracted**	Null	Large pre-registered longitudinal studies tracking actual hormones failed to find shifts in preference for masculine faces/traits across the cycle.

Power Posing (Hormones & Risk)	**Failed / Retracted**	Null	Large pre-registered replications (Ranehill et al. 2015, $N =$) found zero effect on T, cortisol, or risk-taking. Original author disavowed.[79]
Single-Sex Education Academic Benefit	**Failed**	Trivial ($g \leq$)	Meta-analysis of 1.6M students shows zero to trivial academic benefit when adequately controlling for prior achievement and SES.[75]
fMRI Sex-Difference Claims	**Failed**	Null	Early literature characterized by small N and massive false-positive rates; most highly localized functional claims have failed in mega-analyses.[5]
'Women Speak More'	**Failed**	Null	Meta-analyses utilizing electronic EAR recording devices show men and women speak roughly the same number of words per day ($\sim$).
Stereotype-Threat	**Failed (in some**	Null to Small	While real in

Effects	**domains)**		specific contexts, the claim that stereotype threat explains the entirety of the math/STEM gap has failed in large-scale replications.

10-Line Summary of Robust Findings

1. Males exhibit substantially greater absolute upper-body physical strength than females, a physiological dimorphism that persists even when controlling for lean body mass.
2. The Greater Male Variability Hypothesis is strongly supported, showing that males exhibit wider statistical variance in brain morphology, cognitive performance, and physical traits.
3. Men reliably experience a significant, adaptive drop in testosterone upon becoming partnered, engaged fathers, shifting physiology from mating to parenting effort.
4. Structural neuroanatomy reveals robust average differences, with males possessing larger total brain volumes and distinct gray matter distributions, though multivariate classification methodologies remain debated.
5. Cross-culturally, there is a massive, highly replicable divergence in vocational interests, with men strongly preferring to work with things and women strongly preferring to work with people.
6. Evolutionary mate preferences hold across cultures: men universally prioritize cues of youth and physical attractiveness, while women prioritize cues of resource acquisition and maturity.
7. While single personality traits show only small univariate sex differences (supporting the gender similarities hypothesis), aggregating these traits into a multivariate profile reveals highly distinct, non-overlapping male and female personality distributions.
8. The gender-equality paradox is empirically real; as nations become wealthier and more egalitarian, sex differences in STEM participation and personality preferences paradoxically widen.
9. Empathy and theory-of-mind tests show a female advantage, but this gap vanishes when men are offered financial incentives, suggesting differences in motivation rather than innate capability.
10. The role of oxytocin has been rigorously redefined; it is not a universal prosocial "cuddle" hormone, but rather a driver of parochial altruism that enhances in-group cooperation while facilitating out-group hostility.

Contested or Open Questions

1. **The Mechanism of the Gender-Equality Paradox:** Do egalitarian societies widen sex differences by allowing intrinsic, biological preferences to manifest freely (the resource hypothesis), or do wealthy societies generate new, horizontal essentialist stereotypes that

culturally sort men and women?
2. **The Nature of the Brain:** Does the human brain exhibit true sexual dimorphism (as argued by multivariate statistical models), or is it a "mosaic" of overlapping traits where binary categorization is merely a statistical artifact?
3. **Endogenous Sex Drive:** Is the higher reported male sex drive biologically anchored, or primarily the result of distinct gendered socialization, sexual scripts, and measurement biases exposed by physiological and "bogus pipeline" data?
4. **Motivation vs. Ability in Cognition:** To what extent are apparent male deficits in empathy, and female deficits in specific visuospatial tasks, artifacts of testing environments, or differential motivation?

Most Important Sources to Read in Full

1. **Hyde, J. S. (2005).** *The Gender Similarities Hypothesis.* American Psychologist. (The foundational text for univariate psychological similarities).
2. **Del Giudice, M., Booth, T., & Irwing, P. (2012).** *The Distance Between Mars and Venus: Measuring Global Sex Differences in Personality.* PLoS ONE. (The definitive counter-argument using multivariate statistics).
3. **Su, R., Rounds, J., & Armstrong, P. I. (2009).** *Men and Things, Women and People: A Meta-Analysis of Sex Differences in Interests.* Psychological Bulletin. (Comprehensive analysis of vocational sorting).
4. **Walter, K. V., et al. (2020).** *Sex Differences in Mate Preferences Across 45 Countries: A Large-Scale Replication.* Psychological Science. (State-of-the-art replication of evolutionary mating psychology).
5. **Stoet, G., & Geary, D. C. (2018).** *The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education.* Psychological Science. (Key evidence that egalitarianism widens STEM gaps).
6. **Breda, T., et al. (2020).** *Gender Stereotypes Can Explain the Gender-Equality Paradox.* PNAS. (The primary counter-thesis to Stoet & Geary).
7. **Ruigrok, A. N., et al. (2014).** *A Meta-Analysis of Sex Differences in Human Brain Structure.* Neuroscience & Biobehavioral Reviews. (The most robust mapping of neuroanatomical differences).
8. **Pahlke, E., Hyde, J. S., & Allison, C. M. (2014).** *The Effects of Single-Sex Compared With Coeducational Schooling on Students' Performance and Attitudes: A Meta-Analysis.* Psychological Bulletin. (Debunking the academic efficacy of single-sex education).
9. **Gettler, L. T., et al. (2011).** *Longitudinal Evidence That Fatherhood Decreases Testosterone in Human Males.* PNAS. (Landmark study on paternal endocrinology).
10. **De Dreu, C. K. W., et al. (2010).** *The Neuropeptide Oxytocin Regulates Parochial Altruism in Intergroup Conflict Among Humans.* Science. (Pivotal reframing of oxytocin from a prosocial agent to an ethnocentric modulator).

Implications for Living Together

When organizing relationships, families, and societies, recognizing the overlapping statistical nature of sex differences is paramount. Navigating these differences requires strictly avoiding

the naturalistic fallacy: the observation that a trait or behavior is biologically rooted or statistically prevalent does not mean it is morally good, inevitable, or appropriate for modern social architecture.

For couples and families, recognizing the flexibility of human endocrinology—most notably the robust paternal testosterone drop—highlights that men are biologically equipped for intensive, nurturing caregiving. This undermines rigid historical divisions of labor that rely on biological essentialism to exclude men from the domestic sphere.

For organizations and societies, understanding the "Things vs. People" interest divergence helps explain occupational clustering in STEM and caregiving fields without necessarily defaulting to the assumption of systemic discrimination. However, because these are statistical averages, institutional structures must remain strictly meritocratic and agnostic to sex. A multivariate difference between populations still guarantees that millions of individuals will possess talents lying entirely outside their sex's statistical mean. Societies and small groups thrive best when policies accommodate broad, population-level trends to maximize overall well-being, without ever using those trends to gatekeep, stereotype, or constrain the potential of the individual.

Works cited

1. A Meta-Analysis of Sex Differences in Physical Ability - ResearchGate, https://www.researchgate.net/publication/237015302_A_Meta-Analysis_of_Sex_Differences_in_Physical_Ability_Revised_Estimates_and_Strategies_for_Reducing_Differences_in_Selection_Contexts
2. Effects of Physical Exercise on Women with Disabilities in South Korea, https://www.mdpi.com/1660-4601/18/23/12791
3. Sex differences in wrist strength: a systematic review - PMC, https://pmc.ncbi.nlm.nih.gov/articles/PMC10725665/
4. Sex Differences in Fearful Personality Traits Are Mediated by, https://www.pdurkee.com/files/pubs/Mansonetal_2022.pdf
5. Sex Differences in the Adult Human Brain: Evidence from 5216 UK, https://pmc.ncbi.nlm.nih.gov/articles/PMC6041980/
6. SEX DIFFERENCES IN THE HUMAN BRAIN 1 - bioRxiv preprint, https://www.biorxiv.org/content/10.1101/123729v3.full.pdf
7. A meta-analysis of sex differences in human brain structure - PMC, https://pmc.ncbi.nlm.nih.gov/articles/PMC3969295/
8. Differences Between Male and Female Brains, https://psychosomatic-osteopathy.com/en/differences-between-male-and-female-brains/
9. Sex and mental health are related to subcortical brain microstructure, https://escholarship.org/content/qt5xk2697n/qt5xk2697n.pdf
10. Joel et al.'s method systematically fails to detect large, consistent sex, https://marcodg.net/wp-content/uploads/2016/03/delgiudice_etal_2016_letter_joel-etal_pnas.pdf
11. Scientific American - September 2017,

https://gwen.johnshopkins.edu/Files/ScientificAmerican_201709.pdf
12. Sex beyond the genitalia: The human brain mosaic - PNAS, https://www.pnas.org/doi/10.1073/pnas.1509654112
13. Sex differences and the neurobiology of affective disorders - PMC, https://pmc.ncbi.nlm.nih.gov/articles/PMC6235863/
14. Sharing and caring: Testosterone, fathering, and generosity among, https://pmc.ncbi.nlm.nih.gov/articles/PMC7508877/
15. Does Cosleeping Contribute to Lower Testosterone Levels in, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0041559
16. Evidence for an adolescent sensitive period to family experiences, https://www.pnas.org/doi/10.1073/pnas.2202874119
17. Men's status and reproductive success in 33 nonindustrial societies, https://www.pnas.org/doi/10.1073/pnas.1606800113
18. Testicular volume is inversely correlated with nurturing-related brain, https://www.pnas.org/doi/10.1073/pnas.1305579110
19. Longitudinal evidence that fatherhood decreases testosterone in, https://pmc.ncbi.nlm.nih.gov/articles/PMC3182719/
20. The descent of a man's testosterone - PNAS, https://www.pnas.org/doi/10.1073/pnas.1113323108
21. Gender Differences and Variability in Creative Ability: A Systematic, https://www.researchgate.net/publication/374496341_Gender_differences_and_variability_in_creative_ability_A_systematic_review_and_meta-analysis_of_the_greater_male_variability_hypothesis_in_creativity
22. Janet Shibley Hyde Sinks Stereotypes With Data, https://www.psychologicalscience.org/observer/janet-shibley-hyde-sinks-stereotypes-with-data
23. The Gender Similarities Hypothesis | Request PDF - ResearchGate, https://www.researchgate.net/publication/7589797_The_Gender_Similarities_Hypothesis
24. The Gender Similarities Hypothesis, https://www.apa.org/pubs/journals/releases/amp-606581.pdf
25. Understanding the Magnitude of Psychological Differences Between, https://pmc.ncbi.nlm.nih.gov/articles/PMC9442632/
26. Men and Women: No Big Difference, https://www.apa.org/topics/personality/men-women-difference
27. Measuring Global Sex Differences in Personality | PLOS One, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0029265
28. measuring global sex differences in personality - PubMed, https://pubmed.ncbi.nlm.nih.gov/22238596/
29. The Distance Between Mars and Venus: Measuring Global Sex, https://pmc.ncbi.nlm.nih.gov/articles/PMC3251566/
30. Multivariate Versus Univariate Conceptions of Sex Differences, https://everyone.plos.org/2012/01/04/multivariate-versus-univariate-conceptions-of-sex-differences-let-the-contest-begin/
31. Self-reported binary gender prediction from personality traits,

https://journals.plos.org/plosone/article/file?type=printable&id=10.1371/journal.pone.0355193
32. (PDF) Men and Things, Women and People: A Meta-Analysis of Sex, https://www.researchgate.net/publication/38061313_Men_and_Things_Women_and_People_A_Meta-Analysis_of_Sex_Differences_in_Interests
33. a meta-analysis of sex differences in interests - PubMed, https://pubmed.ncbi.nlm.nih.gov/19883140/
34. All STEM fields are not created equal: People and things interests, https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2015.00189/full
35. Do Interests and Cognitive Abilities Help Explain College Major, https://d-nb.info/1209879476/34
36. Gender stereotypes about psychological science, https://epub.uni-regensburg.de/77670/1/1-s2.0-S0001691825007565-main.pdf
37. Sex Differences in Mate Preferences | PDF - Scribd, https://www.scribd.com/document/822553446/walter-et-al-2020-sex-differences-in-mate-preferences-across-45-countries-a-large-scale-replication
38. Sex Differences in Mate Preferences Across 45 Countries, https://www.psychologicalscience.org/journals/psychological-science/0956797620904154/
39. Sex Differences in Mate Preferences Across 45 Countries, https://labs.la.utexas.edu/buss/files/2020/03/Sex-Differences-in-Mate-Preferences-Across-45-Nations-2020.pdf
40. *Deconfounding Sex and Sex of Partner in Mate-Preference Research, https://pmc.ncbi.nlm.nih.gov/articles/PMC13428870/
41. Why might romantic partner preferences vary across cultures., https://www.cambridge.org/core/blog/2026/09/10/why-might-romantic-partner-preferences-vary-across-cultures/
42. Testing the Empathizing–Systemizing theory of sex differences and, https://www.pnas.org/doi/10.1073/pnas.1811032115
43. Testing the Empathizing–Systemizing theory of sex differences and, https://www.researchgate.net/publication/328891377_Testing_the_Empathizing-Systemizing_theory_of_sex_differences_and_the_Extreme_Male_Brain_theory_of_autism_in_half_a_million_people
44. Do Autistic People Have Degrees of Disability in Theory of Mind, https://www.tandfonline.com/doi/full/10.1080/1047840X.2025.2605593
45. Money Affects Theory of Mind Differently by Gender | PLOS One, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0143973
46. Emotions and education: cultivating compassionate minds. PhD thesis., https://theses.gla.ac.uk/3844/1/2012mackenziephd.pdf
47. Libido - Wikipedia, https://en.wikipedia.org/wiki/Libido
48. Is There a Gender Difference in Strength of Sex Drive? Theoretical, https://www.researchgate.net/publication/238048296_Is_There_a_Gender_Difference_in_Strength_of_Sex_Drive_Theoretical_Views_Conceptual_Distinctions_and_a_Review_of_Relevant_Evidence

49. Sexual Dials (Not Switches) Theory: An Evolutionary Perspective on, https://cdn2.psychologytoday.com/assets/2023-08/Sexual_Dials_Not_Switches_Theo.PDF
50. A 48-nation study of sex, culture, and strategies of human mating, https://www.researchgate.net/publication/7564691_Sociosexuality_from_Argentina_to_Zimbabwe_A_48-nation_study_of_sex_culture_and_strategies_of_human_mating
51. (PDF) Beyond Global Sociosexual Orientations - ResearchGate, https://www.researchgate.net/publication/23422008_Beyond_Global_Sociosexual_Orientations_A_More_Differentiated_Look_at_Sociosexuality_and_Its_Effects_on_Courtship_and_Romantic_Relationships
52. Salivary oxytocin has nonlinear relationships with trust and reciprocity, https://www.econstor.eu/bitstream/10419/330010/1/games-14-00017.pdf
53. Oxytocin promotes altruistic punishment - PMC - NIH, https://pmc.ncbi.nlm.nih.gov/articles/PMC5714236/
54. Oxytocin promotes human ethnocentrism - PNAS, https://www.pnas.org/doi/10.1073/pnas.1015316108
55. Oxytocin promotes coordinated out-group attack during intergroup, https://elifesciences.org/articles/40698
56. Oxytocin promotes coordinated out- group attack during intergroup, http://www.joerg-gross.net/files/publications/article/2019_oxytocin-promotes-coordinated-out-group-attack.pdf
57. Variability hypothesis - Wikipedia, https://en.wikipedia.org/wiki/Variability_hypothesis
58. A Key Characteristic of Sex Differences in the Developing Brain - PMC, https://pmc.ncbi.nlm.nih.gov/articles/PMC6041809/
59. Sex Differences in Variability in General Intelligence: A New Look at, https://www.researchgate.net/publication/240323443_Sex_Differences_in_Variability_in_General_Intelligence_A_New_Look_at_the_Old_Question
60. Gender Differences in the Distribution of Creativity Scores - Frontiers, https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2021.626911/full
61. arXiv:1703.04184v15 [q-bio.PE] 2 Oct 2024, https://arxiv.org/pdf/1703.04184
62. An Evolutionary Theory for the Variability Hypothesis - arXiv, https://arxiv.org/html/1703.04184v15
63. (PDF) The Gender-Equality Paradox in Science, Technology, https://www.researchgate.net/publication/323197652_The_Gender-Equality_Paradox_in_Science_Technology_Engineering_and_Mathematics_Education
64. The Gender-Equality Paradox in Science, Technology, Engineering, https://fibamdp.wordpress.com/wp-content/uploads/2018/03/the-gender-equality-paradox.pdf
65. Gender-equality paradox - Wikipedia, https://en.wikipedia.org/wiki/Gender-equality_paradox
66. The Gender-Equality Paradox in Science, Technology, Engineering, https://pubmed.ncbi.nlm.nih.gov/29442575/

67. Relationship of gender differences in preferences to economic, https://www.researchgate.net/publication/328370172_Relationship_of_gender_differences_in_preferences_to_economic_development_and_gender_equality
68. Gender gaps in patience, risk-taking, trust, and prosociality are, https://academic.oup.com/pnasnexus/article/5/8/pgag255/8758741
69. Gender stereotypes can explain the gender-equality paradox, https://dauphine.psl.eu/en/women-and-science/publications/gender-stereotypes-can-explain-the-gender-equality-paradox
70. Gender stereotypes can explain the gender-equality paradox - PNAS, https://www.pnas.org/doi/10.1073/pnas.2008704117
71. (PDF) Gender Stereotypes Can Explain the Gender-Equality Paradox, https://www.researchgate.net/publication/386358355_Gender_Stereotypes_Can_Explain_the_Gender-Equality_Paradox
72. Gender stereotypes can explain the gender-equality paradox - PNAS, https://www.pnas.org/doi/abs/10.1073/pnas.2008704117
73. Simpson's gender-equality paradox - PNAS, https://www.pnas.org/doi/pdf/10.1073/pnas.2422247122
74. Is Single-Sex Education Beneficial for Child Learning? - Consensus AI, https://consensus.app/home/blog/is-single-sex-education-beneficial-for-child-learning/
75. The Effects of Single-Sex Compared With Coeducational Schooling, https://www.researchgate.net/publication/260089187_The_Effects_of_Single-Sex_Compared_With_Coeducational_Schooling_on_Students'_Performance_and_Attitudes_A_Meta-Analysis
76. The Effects of Single-Sex Compared With Coeducational Schooling, https://www.apa.org/pubs/journals/releases/bul-a0035740.pdf
77. Copy of Promises and Pitfalls of Single-Sex Education: Final Report, https://aibm.org/wp-content/uploads/2025/07/Promises-and-Pitfalls-of-Single-Sex-Education_-Final-Report_For-Public-Release.pdf
78. Single‐sex schooling, gender and educational performance, https://researchrepository.ul.ie/bitstreams/76e70e35-0473-49ca-ad0f-44392d4d6182/download
79. Assessing the Robustness of Power Posing: No Effect on Hormones, https://faculty.haas.berkeley.edu/dana_carney/ranehill.pdf
80. Assessing the Robustness of Power Posing: No Effect on Hormones, https://www.researchgate.net/publication/274093025_Assessing_the_Robustness_of_Power_Posing_No_Effect_on_Hormones_and_Risk_Tolerance_in_a_Large_Sample_of_Men_and_Women
81. The effects of power posing on neuroendocrine levels and risk-taking, https://pmc.ncbi.nlm.nih.gov/articles/PMC11619626/
82. Does Power Posing Alter the Intrapsychic Experience of Power?, https://spsp.org/news-center/character-context-blog/does-power-posing-alter-intrapsychic-experience-power
83. CRSP special issue on power poses: what was the point and what, https://www.researchgate.net/publication/317957987_CRSP_special_issue_on_po

wer_poses_what_was_the_point_and_what_did_we_learn
84. Power Posing | Encyclopedia MDPI, https://encyclopedia.pub/entry/37582