


COMMENTARY

Open Access



Sex differences in the human brain: a roadmap for more careful analysis and interpretation of a biological reality

Alex R. DeCasien^{*} , Elisa Guma, Siyuan Liu and Armin Raznahan

Abstract

The presence, magnitude, and significance of sex differences in the human brain are hotly debated topics in the scientific community and popular media. This debate is largely fueled by studies containing strong, opposing conclusions: either little to no evidence exists for sex differences in human neuroanatomy, or there are small-to-moderate differences in the size of certain brain regions that are highly reproducible across cohorts (even after controlling for sex differences in average brain size). Our *Commentary* uses the specific comparison between two recent large-scale studies that adopt these opposing views—namely the review by Eliot and colleagues (2021) and the direct analysis of ~40k brains by Williams and colleagues (2021)—in an effort to clarify this controversy and provide a framework for conducting this research. First, we review observations that motivate research on sex differences in human neuroanatomy, including potential causes (evolutionary, genetic, and environmental) and effects (epidemiological and clinical evidence for sex-biased brain disorders). We also summarize methodological and empirical support for using structural MRI to investigate such patterns. Next, we outline how researchers focused on sex differences can better specify their study design (e.g., how sex was defined, if and how brain size was adjusted for) and results (by e.g., distinguishing sexual dimorphisms from sex differences). We then compare the different approaches available for studying sex differences across a large number of individuals: direct analysis, meta-analysis, and review. We stress that reviews do not account for methodological differences across studies, and that this variation explains many of the apparent inconsistencies reported throughout recent reviews (including the work by Eliot and colleagues). For instance, we show that amygdala volume is consistently reported as male-biased in studies with sufficient sample sizes and appropriate methods for brain size correction. In fact, comparing the results from multiple large direct analyses highlights small, highly reproducible sex differences in the volume of many brain regions (controlling for brain size). Finally, we describe best practices for the presentation and interpretation of these findings. Care in interpretation is important for all domains of science, but especially so for research on sex differences in the human brain, given the existence of broad societal gender-biases and a history of biological data being used justify sexist ideas. As such, we urge researchers to discuss their results from simultaneously scientific and anti-sexist viewpoints.

*Correspondence: decasienar@nih.gov

Section On Developmental Neurogenomics, National Institute of Mental Health, Bethesda, MD, USA



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Highlights

- Recent large-scale studies have reached different conclusions regarding the presence of sex differences in human neuroanatomy.
- We show that these contradictory findings are explained by different methodological choices. While multiple large direct analyses highlight small, highly reproducible sex differences, reviews do not account for methodological heterogeneity across studies (e.g., statistical power/sample size, brain size-correction methods, segmentation, region selection, participant age). This explains many of the apparent inconsistencies reported in recent reviews.
- We also summarize observations that motivate research on sex differences in human neuroanatomy (including potential causes and effects), review methodological and empirical support for using structural MRI to investigate such patterns, and outline best practices for analyzing and describing neuroanatomical sex differences.
- Finally, we argue that broader historical and societal contexts make it important to reinforce the scientific method by adopting an actively "anti-sexist" viewpoint when conducting research on sex differences in the human brain.

Keywords: Sex differences, Neuroanatomy, sMRI, Direct analysis, Meta-analysis, Review, Sex chromosomes, Neurodevelopment, Sexual selection, Anti-sexism

Background

Introduction

The presence and magnitude of sex differences in human neuroanatomy are hotly debated topics in neuroscience [1–5]. While some studies argue there is little to no evidence for notable, consistent sex differences in human brain structure [1, 6, 7], others report highly reproducible, small-to-moderate sex differences in the size of multiple brain areas, even after controlling for differences in average brain size [2, 3, 8, 9]. With a mounting wealth of data at hand [10], and evermore powerful tools for data analysis [11], how can such contrary beliefs persist around a seemingly simple empirical question? This discrepancy is brought to a head by the titles of two recent papers that spur our current *Commentary*: (1) “*Dump the ‘dimorphism’: Comprehensive synthesis of human brain studies reveals few male–female differences beyond size*” by Eliot and colleagues [1]—a “meta-synthesis” of previously published structural magnetic resonance imaging (sMRI) studies on neuroanatomical sex differences; and (2) “*Sex differences in the brain are not reduced to differences in body size*” by Williams and colleagues [3]—a response to the former article that describes a direct analysis of sMRI data in ~40k individuals (aged 40–70 years) [2] from the UK Biobank neuroimaging dataset [12]. Is it possible to reconcile these starkly contrasting conclusions and explain how such conflicting beliefs could emerge?

Our *Commentary* uses the specific comparison between these two recent large-scale studies [1, 3] as a path towards: (i) clarifying the controversy surrounding the topic of sex differences in human neuroanatomy; and

(ii) providing a structured framework for the planning, execution, and interpretation of this research. Our goal is to limit miscommunication and promote more effective progress in both scientific and lay discourses on sex differences in human brain anatomy. Throughout this *Commentary*, we make the case that research on sex differences should apply the scientific method in an unwaveringly principled and impartial way, while also actively combating the risks of false extrapolations that draw unfounded conclusions with no support from the data. This approach precludes research that is motivated either by sexism or by the desire to reject the existence of any biological sex differences. There is an undeniable history of biological data being used in attempts to justify sexist views [13–15], and while reporting neuroanatomical sex differences is not an inherently sexist act, modern research on this topic is not isolated from its history. Thankfully, the scientific method offers a clear and objective scaffolding that we can use to help us safely engage with crucial questions [16] regarding potential sex differences in the human brain.

The sections below are organized according to the phases of the scientific method. First, we consider observations that motivate research on sex differences in human brain anatomy. We organize these observations into four groups: (1) concepts from evolutionary theory (i.e., potential distal causes of sex differences in human brain organization); (2) genetic and environmental factors that differ between males and females (i.e., potential mechanistic causes of sex differences in human brain organization); (3) sex differences in the prevalence and presentation of brain-based disorders

(i.e., epidemiological and clinical evidence for potential sex differences in human brain organization); and (4) the specific case for using in vivo sMRI as a tool for probing sex differences in human brain organization. Next, we outline best practices for investigations of sex differences in human brain structure, organized according to the following three sequential steps of the scientific method: specifying the question, designing the appropriate study, and framing and interpreting the results. In considering each of these research steps, we highlight how the choices an investigator makes can profoundly impact their final conclusions regarding the existence and significance of sex differences in human brain anatomy. This process helps to make sense of the contrasting results reported by the two recent studies that motivate our *Commentary* [1, 3]. It also offers a more general road map that will hopefully aid effective progress in research on sex differences in the human brain.

In the following sections, we use the term biological “sex” to refer to an individual’s combined sex chromosome complement and gonadal phenotype, which together group most individuals into two categories: males as XY individuals with testes and females as XX individuals with ovaries. Although most individuals can be categorized as female or male based on this system (a designation which is typically made based on genital phenotype at birth), sex is not strictly binary. In fact, intersex individuals, who represent ~1% of the population, exhibit an abundance of variation across sex chromosome combinations, sex hormone concentrations, and bodily phenotypes. Bearing in mind this variation, researchers may still meaningfully use biological sex as a concept to examine group-level differences between individuals with “typical” (of the majority) anatomical and sex chromosome combinations, although we acknowledge that these criteria are not confirmed in many of the human studies discussed here, which instead rely on self-identification. Specifically—as for all statistical analyses in biological research—any true group differences should be detectable with sufficient sample sizes, even if there is some amount of intra-group variability or misclassification. Finally, we distinguish biological sex from gender, a culturally defined and malleable concept, and recognize that an individual’s gender need not align with their biological sex. Given that an individual’s experiences in society can be impacted by their perceived gender, biological and psychosocial influences are often bidirectionally entangled in humans, making their effects difficult to untwine [17]. Amidst this complexity, characterizing phenotypic differences based on biological sex provides one empirically tractable pathway towards better understanding the intertwined influences of sex and gender on the human brain.

Main text

Part 1: Observations that motivate research on sex differences in regional brain anatomy

The idea that humans may show sex differences in brain organization is prompted by a wealth of prior theoretical and empirical information from evolutionary, neuroscientific, and medical research. Evolution provides potential distal causes for *why* a phenotypic sex difference may exist (e.g., sex differences in body size as a result of competition over mates), whereas genetic and environmental factors provide potential mechanistic causes that address *how* a sex difference emerges during development (e.g., sex differences in body size being achieved through the sex-biased action of gonadal steroids) [18, 19]. These distinct causal frameworks are complementary (yet sometimes interwoven: [20]), such that attaining an understanding of a mechanistic cause does not preclude the existence of evolutionary explanation [21]. Beyond consideration of causal factors, the high a priori likelihood of human showing sex differences in brain organization is also supported by the existence of very many well-documented sex differences in prevalence, presentation, and prognosis of diverse medical disorders involving the brain. We expand upon each of these motivations for studying sex differences in human brain organization below, before considering why we might expect some of these differences to manifest as sex differences in regional brain anatomy that are resolvable by in vivo sMRI.

Potential distal causes of sex differences in human brain organization

150 years ago, Charles Darwin proposed the concept of sexual selection to explain the numerous sex differences (both physiological and behavioral) that he observed throughout the animal kingdom [22]. He outlined two mechanisms, including: (i) mate choice—individuals of sex A who possess a certain version of a trait are more attractive to (and likely to be chosen as a mate by) members of sex B; and (ii) mate competition—individuals of sex A who possess a certain version of a trait can out-compete other members of sex A for mating opportunities with members of sex B. In either case, the individuals exhibiting these trait varieties produce more offspring, leading them to become more prevalent over generations in sex A but not in sex B (to whom these varieties are not beneficial). These mechanisms have not only produced sex differences in visible, physical traits across the animal kingdom—such as the elaborate tails of male peacocks and the large antlers of male deer—but they have also impacted brain evolution. For example, mate choice is responsible for the elegant songs (and larger brain song nuclei) among the males of many songbird species [23], while competition over locating mates has led to greater

spatial navigation skills (and larger hippocampi) among the males of certain vole species [24].

Like all other animals, the human species emerged as a result of evolutionary processes that shaped our bodies and behaviors. Although we are distinct from nonhuman animals in many ways (e.g., theory of mind, language, complex tool use) [25], we are not immune to our own evolutionary history, one that certainly included some aspects of sexual selection. For example, when we look across primate species, mate competition tends to be stronger in species with larger sex differences in body size [26]. Accordingly, an evolutionary history involving competition over mates (among males) is among the most viable evolutionary explanations for sex differences in human height and weight [26]. While hypothetical evolutionary causes are hard to empirically assess, sexual selection has provided a powerful explanatory framework for additional sex-biased traits in humans, including facial hair [27], fat distribution [28], aggression [29], and spatial ability [30]. For example: i) mate competition among males may contribute to male-biased rates of aggression across human cultures [29]; ii) sex differences in foraging throughout early human evolution (and after transitions to agriculture in some populations [31]) may partially explain sex-biased spatial abilities (e.g., labor division: females gather from spatially stable but seasonally variable food sites; males hunt across long distances spanning various routes; modern performance patterns: females outperform on object location memory and navigation by landmark tasks; males outperform on mental rotation tasks, which are associated with throwing accuracy and navigation by orientation) [32, 33]; and (iii) effective mate choice may require behavioral inhibition and contribute to female outperformance on inhibitory tasks [34]. If sexual selection did, in fact, shape adult human behavior over evolutionary time, then this will have necessarily been achieved through sex differences in neurodevelopment and adult brain organization. As discussed below, the leading proximal sources of sex-biased brain development are sex differences in chromosomal complement and gonadal type.

Potential mechanistic causes of sex differences in human brain organization

The hypothesis that humans may show sex differences in brain organization is not only raised by evolutionary theory, but also by genetic and environmental differences between males and females. Individual differences in sex chromosome complement and the concentration of gonadal sex steroids are foundational to the biological definition of sex in humans and nonhuman mammals, and there is substantial experimental evidence from animal studies that these factors directly impact mammalian

brain organization [35]. Accordingly, it is reasonable to expect that human brain development is also subject to direct gonadal and sex chromosome effects. As detailed further below, several lines of observational data support the existence of such effects in humans.

In placental (eutherian) mammals, presence of the *SRY* gene on the Y chromosome of XY males typically leads to differentiation of the embryonic gonadal ridge into testes, whereas the absence of this gene in XX females typically allows differentiation into ovaries [36, 37]. This gonadal divergence tends to lead males and females to exhibit distinct concentration profiles of various sex steroids (e.g., estrogens, androgens) throughout development, although these hormones are not limited to one sex and their concentration distributions overlap [38]. Sex hormones were first shown to shape sex-biased mammalian (rodent) brain organization in the late 1950's [39], and these findings have been bolstered and extended by the application of a growing suite of experimental methods in murine research [35, 40, 41]. Some of these gonadal influences actually exert experimentally verifiable influences on murine brain anatomy as measured by sMRI [42, 43]. Evidence for similar links between sex steroids and neuroanatomy in humans is largely derived from observational neuroimaging studies that harness non-experimental variations in human sex steroid signaling as a function of: (i) inter-individual variation in circulating sex steroids during development, reflecting variation between individuals in the same developmental period and also across developmental periods (e.g., pre/post puberty, pre/post menopause) [44, 45]; (ii) the menstrual cycle [46]; (iii) medical disorders impacting the hypothalamo–pituitary–gonadal axis [47]; or (iv) gender-affirming hormone treatment [48].

Although gonadal steroids have been historically viewed as the primary mechanistic drivers of sex differences in mammalian brain organization, there is growing evidence that sex chromosome complement can also have direct effects on mammalian brain organization [49, 50]. For instance, the products of sex chromosome genes are present at different levels in male and female cells [50]: genes on the Y chromosome are expressed only in males and genes that escape X chromosome inactivation (XCI) are often expressed at higher levels females [51, 52]. In addition, XCI in females may partially monopolize cellular epigenetic machinery, altering the expression of other genes [53]. Thus, irrespective of gonadal differences, XY and XX individuals have constitutional differences in the dosage of genes that are known to exert genome-wide regulatory effects that are not mediated by the gonads [54, 55]. This is bolstered by reports that sex-biased expression of autosomal genes in the mouse brain occurs prior to gonadal differentiation [56]. Moreover,

there is extensive evidence in both rodents and humans that sex chromosome dosage can influence brain organization [50]. The four core genotype (FCG) model in mice provides experimental evidence that XX and XY groups show several reproducible neuroanatomical differences in sMRI [42, 43], as well differences in vasopressin fiber density within the lateral septum [57]. Evidence for sex chromosome dosage effects on the human brain comes from observational research in groups with differing X and/or Y chromosome doses due to sex chromosome aneuploidy. Specifically, increasing X- and/or Y- chromosome dosage induces highly reproducible changes in diverse aspects of regional brain anatomy, including parieto-occipital cortical surface area (increased), lateral temporal cortical thickness (decreased), and volume of the cerebellum and globus pallidus (both decreased) [58–61].

Alongside the aforementioned effects of gonadal and sex chromosome dosage on the brain, there are also extensively well-documented physiological sex differences in multiple peripheral organs and tissues that may affect the brain. These include adipose [62], hepatic [63], peripheral immune [64], and renal systems [65]. Such somatic sex differences are notable on two grounds: i) they are far less debated or controversial than the existence of sex differences in the brain; and ii) they represent another potential source of sex differences in the brain since they guide concentrations of various peptide and steroid hormones (e.g., leptin, adiponectin, inflammatory cytokines, neurotransmitter precursors, and estrogens) [66].

In addition to the many potential evolutionary, genetic, and endocrine-related causes of sex differences in brain organization, males and females are (at the group-level) often exposed to systematically different environments across the lifespan. For example, the experience of pregnancy has been linked to changes in brain organization [67, 68], and XY males do not experience pregnancy (although fatherhood also impacts the brain: [69]). There is also extensive evidence that individuals who are outwardly perceived as male or female experience pervasive differences in diverse domains of life [70], spanning the educational [71], professional [72, 73], fiscal [74], and medical [75] arenas, in addition to risk of exposure to different dangerous situations [76]. These considerable, consistent, and enduring gendered socio-environmental factors could conceivably influence brain organization in a manner that would manifest as group differences between males and females. It can, however, be extremely difficult (or impossible) to experimentally verify the action of such experiential sex differences on human brain organization, and it is likely there may well be synergy between experientially and biologically driven

sex differences. In particular, gendered experiential influences on the human brain may be intimately coupled with sex-biased genetic and endocrine influences [77] since gendered experiences are correlated with an individual's perceived gender. For example, observed sex differences in aggression and spatial ability may reflect, or be amplified by, gender-biases in the acceptance of and encouragement towards certain types of behaviors and toys throughout development [30, 78].

Taken together, the lines of research reviewed above identify strong evolutionary priors for the existence of sex differences in human brain organization, and also specify numerous endocrine and genetic factors that differ between typically developing human males and females and are known to influence mammalian neurodevelopment. When considering these potential sources of sex differences, however, it is also important to consider how they inform our understanding of the many phenotypic *equivalences* between males and females. At one level, the gonadal and chromosomal differences between typical males and females mean that any features of the mammalian brain that lack a sex difference are necessarily being attained from different starting points in each sex. For example, how do two instances of the same neuronal subtype, one in a male and one in a female, arrive at the same structural and functional profile despite having categorically different sex chromosome gene dosages? Similar questions can be asked about organ-level phenotypic equivalence: how do these two populations, with differences in brain size (which reflect differences in body size), arrive at equivalent intelligence levels [79]? One possible path to phenotypic equivalence between the sexes may actually be mechanistic counterbalancing of two opposing sex differences (i.e., compensatory mechanisms that serve to *prevent* sex differences in brain function and behavior) [80]. For example, experiments in the FCG murine model have identified neuroanatomical regions where sex chromosome and gonadal complement exert opposite effects on volume. These “ying-yang” effects appear to exist in brain areas that exhibit sex-biased volumes (e.g., volume of the medial amygdala is larger in mice with testes than mice with ovaries, but smaller in XY vs. XX mice), as well as regions without any apparent sex differences in mean volume (e.g., the pallidum, where volumetric differences between XX and XY mice only emerge after removal of the gonads, indicating that gonadal effects are “neutralizing” chromosomal effects) [43]. This latter scenario is a specific example of the more generalizable idea that a given phenotype can be similar between males and females but supported by different underlying mechanisms. Additional examples can be found at various biological levels, including interactions between the epigenome and transcriptome

(e.g., XCI in females, which results in similar levels of expression for X chromosome genes (that do not escape XCI) in males and females) and between neuroendocrine signaling and behavior (e.g., male and female prairie voles exhibit similar levels of parental care, and while this behavior is produced by pregnancy-related hormonal changes in females, it is driven by vasopressin signaling in males) [80]. Thus, the absence of a sex difference at one level of biological organization cannot be assumed to imply the absence of a sex difference at other levels. Moreover, the preponderance of phenotypic *similarities* between males and females poses fascinating biological questions in light of the aforementioned sex differences in chromosomal dosage and gonadal status.

Epidemiological and clinical evidence for potential sex differences in human brain organization

Some of the strongest indirect evidence for biologically driven sex differences in human brain organization is provided by epidemiological and clinical studies that demonstrate large sex differences in the prevalence, presentation, and progression of brain-based disorders [81, 82]. These differences cannot in and of themselves be assumed to solely reflect biologically programmed sex differences in brain organization given the profound gendered influences on health [75], but several observations suggest that biologically grounded factors are likely to be an important contributor. First, multiple neurodevelopmental disorders with onset in early childhood are more prevalent in males than females [81, 83], including autism spectrum disorders (ASD) [84], attention-deficit hyperactivity disorder (ADHD) [85], and early-onset obsessive compulsive disorder (OCD) [86]. In contrast, there is a sharp rise in female-biased risk for depression [87], anxiety [88] and eating disorders [89] in adolescence, and of certain neurodegenerative (e.g., Alzheimer's disease and other dementias) [90] and psychiatric disorders (e.g., late-onset schizophrenia) [91] later in life. The concentrated emergence of sex-biased disorders in windows of particularly dynamic human brain development (i.e., early childhood, adolescent, and post-menopausal years) suggests that sex differences in neurobiology contribute to the emergence of these conditions. Sex biases in sociocultural norms and the applicability of diagnostic criteria are likely to contribute to some of the observed sex biases in disease rates. For example, reduced help-seeking behavior in men may explain their lower treatment rates for depression [92], and current diagnostic guidelines lead clinicians to under-detect ASD in females [93]. However, such biases cannot explain multiple diagnostically distinct disorders all showing the same-sex bias during the same developmental window (e.g., male-biased early-onset ASD, ADHD, and OCD). Second,

several sex-biased psychiatric disorders display biological features that interact with known sex differences in neurobiology. For example, the profile of altered gene expression from postmortem brain studies in ASD is partly correlated with normative sex differences in brain gene expression [94]. Similarly, the adolescent-emergent female bias in anxiety is closely linked to measures of pubertal progression that track activation of the hypothalamo–pituitary–ovarian axis in females [88]. Third, medical disorders that modify sex-linked biological factors (e.g., sex chromosome dosage) are associated with increased risk for sex-biased brain disorders [95]. Fourth, males and females exhibit differences in the presentation and progression of multiple neurological conditions. For example: i) females with ASD exhibit more camouflaging behaviors (i.e., compensating for and masking autistic characteristics) and a higher rate of co-occurring internalizing disorders that may mask autistic symptoms, and these factors contribute to the under-diagnosis of ASD in females [93]; ii) mania is more prevalent among males with bipolar disorder, while depression and comorbid panic disorders, eating disorders, and borderline personality disorder are more prevalent among females with bipolar disorder [91]; and iii) males with schizophrenia tend to exhibit higher rates of substance abuse and lower rates of recovery and remission [91]. These differences suggest that sex-specific biological and experiential factors may interact with disease-related factors to impact condition progression. Finally, sex differences in treatment responsiveness continue to emerge. The mechanisms underlying these differences include sex effects on drug metabolism (e.g., in Parkinson's and schizophrenia treatments) [96, 97] and specific sex–genotype interactions that affect treatment response (e.g., APOE genotype \times sex effects on Alzheimer's treatment outcomes) [98]. Taken together, these diverse sex differences across multiple brain-based disorders provide compelling indirect evidence for sex differences in human brain organization.

The case for probing potential sex differences in human brain organization using in vivo structural neuroimaging

The sections above outline a myriad of evolutionary, genetic, and environmental factors that may produce sex differences in human brain organization, potential evidence for which includes documented sex biases in the prevalence and presentation of multiple brain-based diseases. Taken together, these considerations strongly suggest that humans are likely to show sex differences in brain organization. However, this *Commentary* is specifically focused on the contentious issue of whether humans show sex differences in one particular aspect of brain organization, namely regional brain anatomy as measured by in vivo sMRI [1, 3]. What are

the grounds for using sMRI to test for sex differences in human brain organization? This question is fundamental since the human brain offers a vast phenotypic landscape for the manifestation of sex differences, one that spans multiple spatial (from molecules to visible folds) and temporal (from milliseconds to lifespans) scales. We argue that there are both methodological and empirical reasons to test for hypothesized sex differences in human brain organization using sMRI.

From a methodological perspective, *in vivo* neuroimaging currently offers the only available means of measuring the living human brain in large numbers of individuals. Moreover, *in vivo* imaging data provide a homogeneous, brain-wide screen for potential sex differences in regional brain anatomy: either to recover effects that have been hypothesized a priori, or to discover novel effects [99, 100]. These properties stand in contrast to postmortem approaches for studying the human brain which, while they enable molecular and microscale measurements, typically require examination of pre-selected regions in small numbers of individuals [50]. Thus, *in vivo* neuroimaging provides the best available tool to conduct brain-wide tests for sex differences in human brain organization in sample sizes that limit the risk of Type II error (i.e., falsely concluding there is no phenotypic sex difference when one does exist). This is especially true for sMRI, which exists in extremely large datasets [10] and provides a vast array of different structural brain features (e.g., cortical thickness, area, folding, etc.) that can be similarly measured with high reliability across humans of diverse developmental stages (e.g., in contrast to task-based functional neuroimaging) and across humans and nonhuman animal models [11].

Fortunately, beyond the methodological “pushes” to test for sex differences in human brain organization using sMRI, there are also multiple empirical “pulls” for positing that sex differences in mammalian brain organization could manifest in ways that are detectable by sMRI. First, in species like mice, where both postmortem and *in/ex vivo* sMRI data are available, sMRI is clearly able to reliably recover all classical histological foci of sex-biased regional brain volume, and also discover new foci that have been missed by the regionally targeted nature of most postmortem research [35, 50, 100]. Second, newer tools for brain-wide analyses of cell counts in mice show that many foci with sex-biased cellular compositions are also regions that show sex-biased volumes by sMRI [100, 101]. Third, gonadal and sex chromosome factors that tend to differ between typical males and females have been shown to modulate sMRI measures of regional brain volume using experimental methods in mice [42, 43] and observational data in humans [47, 48, 58–61].

Part 2: Specifying the question: do humans show sex differences in regional brain anatomy by sMRI?

The aforementioned observations strongly motivate empirical tests for regional sex differences in human brain volume using *in vivo* sMRI. However, conducting such a test requires first specifying what one means by “sex” and a “sex difference”. That is, we need to properly define the groups being compared, and the observations that would constitute a group difference. When testing for sex differences in regional brain volume, it is also necessary to specify if and how any adjustment will be made for well-established average sex differences in height and total brain volume. We consider each of these issues below.

Specifying terms of reference when referring to “sex”

The need to define terms of reference in formulating a research question is particularly relevant for the study of sex differences. Even the apparently straightforward notion of biological sex is not a clear binary: sex chromosomes are not sex specific, sex hormones are not sex limited (i.e., androgens and estrogens are not male- and female-specific, respectively), and sex presentation does not perfectly reflect circulating hormones or chromosome complement [102]. Accordingly, when testing for sex differences, it is important to explicitly state how sex is being defined. Definitions may include self-reported sex or karyotypic sex.

Specifying terms of reference when referring to a “sex difference”

It is also important to clarify the terminology being used to describe the type of sex difference being considered. Here, we present a slightly modified version of the framework outlined by McCarthy and colleagues [35]: (1a) “sexual dimorphism” should only be used when a trait has two forms, one more prevalent in males and one more prevalent in females (usually limited to traits directly associated with reproduction); (1b) “sexual polymorphism” should be used when a trait has more than two forms (compared to sexual dimorphism, this term more accurately reflects variation from e.g., sex chromosome aneuploidies, intersex physical phenotypes) [103]; (2) “sex difference” should be used when the trait exists on a continuum, with males and females exhibiting mean or variability differences; and (3) “sex convergence and divergence” should be used when the trait does not exhibit a sex difference but comprised different neural underpinnings. In addition, a useful alternative may be to describe “sex differences” as the result of “sex influences” [104], since this term may help prevent incorrect assumptions of large categorical differences. Clarity in terminology is not only important for accurately

conveying one's own research findings to others, but also for properly interpreting published research and discussing perceived conflicts or controversies in research. The conflation between “sexual dimorphisms” and “sex differences” is particularly problematic, and can lead well-intending researchers to erroneously describe potential sex differences as sexual dimorphisms (e.g., [105, 106]), or readers to mistakenly conclude that reported sex differences are being taken by their authors to imply evidence for a sexual dimorphism [1].

Adjustments for height or total brain size

Another critical clarification when studying sex differences in human neuroanatomy relates to the facts that: (1) average body size and overall brain volume are larger in males than in females [3, 107]; and (2) regional volumes show strong positive correlations with overall brain volume that are region-specific in their magnitude [108]. These observations mean that all absolute regional brain volumes are, on average, larger in males than females, and that certain brain areas tend to comprise a larger or smaller proportion of male versus female brains, depending on whether those areas scale hyper- or hypo-allometrically with brain size, respectively [59, 108–111]. Accordingly, many studies seek to control for sex differences in overall brain size when investigating sex differences in regional anatomy. The neurobiological rationale for focusing on sex differences in regional brain volume stems from the partly localized nature of structure–function mappings in the mammalian brain [112, 113], which imply that any domain-specific functional sex differences (e.g., of reproductive or social behavior) are more likely to be related to network-specific (versus global) sex differences in brain organization. Additionally, there is evidence that species differences in relative, but not absolute, brain region size correlate with species differences in behavior and cognition (e.g., relatively, but not absolutely, larger hippocampus volumes are observed in food caching vs. non-caching bird species) [24]. In specifying one's question regarding the existence of sex differences in regional brain anatomy by sMRI, it is not just important to state *if* brain size is being controlled for, but *how* any control is implemented. This is a critical issue because some popular methods for controlling brain size in analyses of group differences induce predictable biases in analyses of neuroanatomical sex differences (see ‘Designing the appropriate study’ below).

The above considerations lead many contemporary analyses of sex differences in the human brain to be guided by a very specific question: “Do self-identified men and women exhibit statistically significant average differences in brain region anatomy, as detectable by *in vivo* sMRI, after controlling for differences in overall

brain size?” This is the question considered by both of the recently published articles that prompted the current *Commentary* [1, 3]. Although both studies considered the same question, and drew on large bodies of data in search of an answer, they arrived at diametrically opposed answers. We show below that this strange state of affairs is a predictable consequence of differences in study design, and we argue that clarity around the impact of different study design choices is a critical ingredient for rational progress on the question of sex differences in human brain anatomy.

Part 3: Designing the appropriate study to test for sex differences in regional human brain anatomy based on sMRI data

A theoretically definitive study design to examine sex differences in regional human brain anatomy based on sMRI data would be to: (i) use the same MRI machine to gather structural brain scans on all male and female humans (defined by e.g., karyotype and circulating sex hormone concentrations) in a target population of interest ascertained using epidemiological sampling principles (e.g., minimizing sampling biases) [114]; (ii) preprocess all scans using an identical analytic pipeline [115]; (iii) use statistical methods to estimate sex differences in regional brain anatomy after considering sex differences in neuroanatomical variability [116, 117], variables that can influence regional brain anatomy (including total brain volume and age), and potential multicollinearity between variables. This study design is impractical, but provides a theoretical ideal against which alternative study designs can be assessed. To date, research on sex differences in regional brain volume has used one of two main study designs, which are considered in turn below.

Approach 1: direct analysis of sex differences in regional brain anatomy using large sMRI datasets

To date, most researchers studying sex differences in human brain anatomy by sMRI have sought to approximate the idealized study design above by analyzing a subsample of the population to estimate true effects in the full population. This study design is exemplified by the recent studies from Williams and colleagues [2, 3], which provide a direct analysis of regional sex differences in brain volume using the largest single-study dataset available to date: ~40,000 sMRI scans from the UK Biobank dataset [12]. These papers concluded that males and females showed statistically significant differences in 67% (409/620) of all cerebral measures analyzed, even after controlling for brain size, age, and technical effects. Given that the vast sample size of this study is likely to facilitate small group differences reaching statistical significance, the authors quantified the distribution of effect

sizes (i.e., standardized betas) for statistically significant group differences, which ranged from -0.67 to 0.64 (median absolute effect size = 0.13). While most differences were “small” (less than 0.1), 46% of regions had an absolute effect size greater than 0.1 (corresponding to a sex difference in region volume equal to 10% of the standard deviation of log-transformed region volume, holding all other factors equal). For example, the reported standardized sex effect for right hippocampal volume was ~0.1 (0.098), which corresponds to females having ~20mm³ (or ~1.5% of the average) larger volumes compared to males with same brain size, of the same age, and whose data were collected under the same technical parameters.

The work by Williams and colleagues [2, 3] demonstrates many of the advantages in testing for sex differences through direct analysis of large sMRI datasets: (i) there is adequate statistical power to confidently detect even small effect sizes; (ii) the effect sizes can be accurately estimated, enabling researchers to go beyond statistical significance and convey the magnitude of observed group differences; and (iii) exquisite analytic control allows investigators to systematically test if and how the statistical significance and effect size of any observed sex differences are altered by consideration of co-occurring technical, age, or brain size effects. Williams and colleagues pay close attention to the issue of brain size control by first systematically mapping non-linear scaling relationships across the human brain, showing that these scaling laws are almost always identical between males and females (with no sex difference for 99.8% of regions), and testing for sex differences in regional brain anatomy using a log–log regression method that avoids the biased effects caused by other brain size adjustment methods [2] (discussed more below). Notably, the work by Williams and colleagues was able to directly test for such complex phenomena (e.g., potential age modulation of scaling effects) and to, therefore, properly account for any such effects (which were found to be minimal in this study [2] and other large-scale reports [111]).

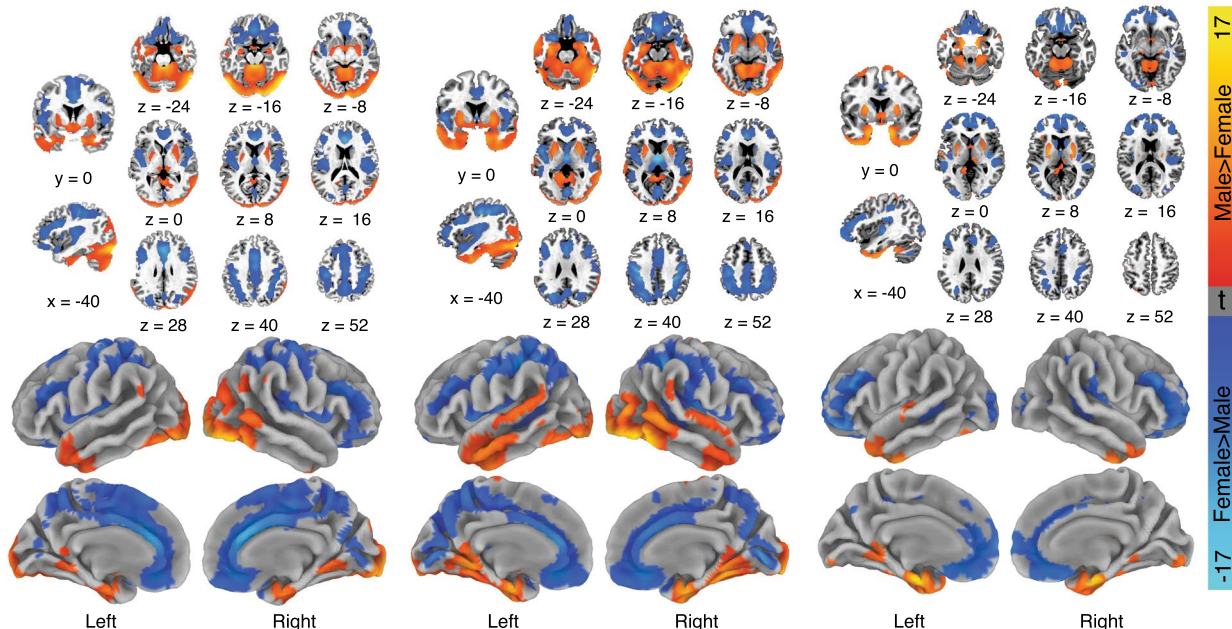
An additional advantage of studying sex differences in human neuroanatomy via direct data analysis was highlighted by two other recent MRI studies [8, 9] which harnessed their large sample sizes to provide direct quantitative tests of the reproducibility of sex differences across different cohorts [8, 9]. Lotze and colleagues [9]

examined two independent cohorts from the Study of Health in Pomerania (SHIP) (N ~ 1000 and N ~ 1800) and found that all regions exhibiting significant sex effects in the smaller cohort were reproduced in the larger cohort. Similarly, Liu and colleagues [8] found that the spatial distribution of sex differences in GMV was highly reproducible across: (i) two large cohorts (N ~ 1000 each) drawn from the Human Connectome Project (HCP) and the UK Biobank (UKB) ($r > 0.80$); and (ii) 1000 split-half samples of the HCP dataset (mean $r \sim 0.8$). Accordingly, these studies demonstrate the reproducibility of sex differences in regional human neuroanatomy by directly comparing estimated sex differences in different samples and quantifying the reproducibility of the map of anatomical sex differences using spatial correlations both within and between samples. Importantly, Liu and colleagues [8] directly quantified the within-cohort reproducibility of sex differences using permutation methods that can accommodate data distributions that violate some assumptions of parametric methods (e.g., the assumption of equal variances in linear regression between males and females). Such focused comparisons of large datasets represent the gold standard approach to reproducibility since, compared to meta-analyses of many methodologically heterogeneous studies (discussed below), they facilitate more straightforward analysis and interpretation. Figure 1 reproduces and directly compares the spatial patterning of significant sex differences in regional gray matter volume (after brain size correction) across the 3 independent samples, including the SHIP cohort (Fig. 1A), HCP cohort (Fig. 1B) and UKB subsample (Fig. 1C). A conjunction of these 3 maps (Fig. 1D) demonstrates that humans show a highly reproducible spatial pattern of sex differences in regional GMV, including male-biased volumes of the putamen, amygdala, hippocampus, and temporal pole and female-biased volumes of the cingulate, superior parietal, and lateral prefrontal cortices. Moreover, this conjunction analysis reveals spatially nested levels of consistency, pointing to steady mounting levels of cross-study agreement approaching specific anatomical foci (Additional file 1: Table S1). Our conjunction also helps to specify the small number of foci with opposing directions of statistically significant sex differences across the 3 studies (Fig. 1E, Additional file 2:

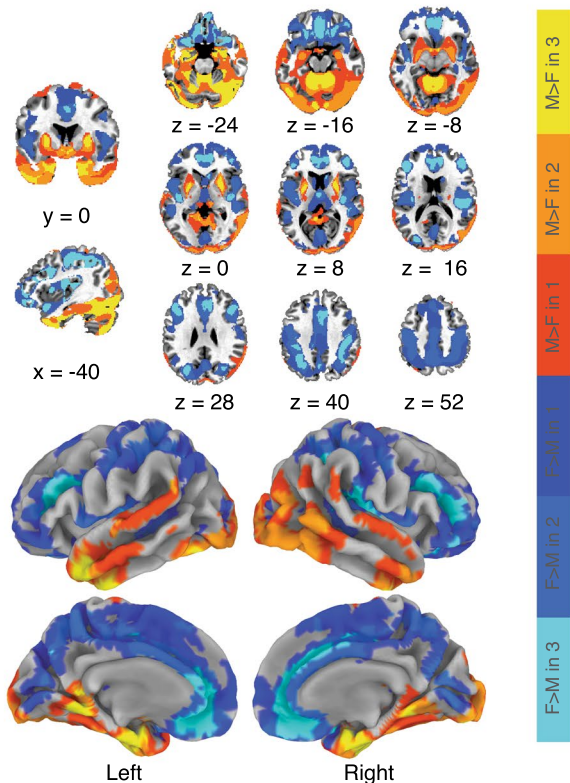
(See figure on next page.)

Fig. 1 The spatial patterning of neuroanatomical sex differences is largely reproducible across 3 large cohorts. Sex differences in regional gray matter volume (after brain size-correction) across 3 independent cohorts. **A** Results from the HCP cohort (N = 976) analyzed by Liu and colleagues [8]. **B** Results from the UKB subsample (N = 1120) analyzed by Liu and colleagues [8]. **C** Results from the SHIP cohort (N = 2838) analyzed by Lotze and colleagues [9]. **D** Conjunction map across 3 cohorts. Color encodes whether 1, 2 or 3 of these cohorts show overlapping statistically significant sex differences in regional GMV (cool colors, F > M, warm colors: M > F). Note the spatial nesting of colors (this is consistent with a core pattern of sex biases in regional GMV that is variably recovered by these three studies). **E** Map of inconsistencies across cohorts. Purple regions are those where any 2 of the 3 cohorts considered showed statistically significant regional GMV sex differences in opposite directions

A) Significant sex differences in HCP (United States; N = 976) **B) Significant sex differences in UKB (United Kingdom; N = 1,120)** **C) Significant sex differences in Lotze (Germany; N = 2,838)**



D) Significant sex differences in the same direction of one or two or all three contrasts



E) Significant sex differences in the opposite direction of any two contrasts

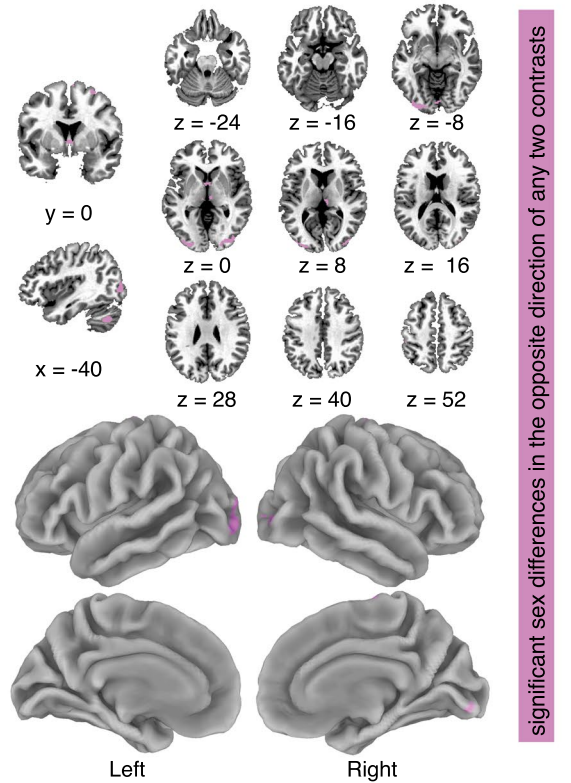


Fig. 1 (See legend on previous page.)

Table S2). We consider how to better understand these inconsistencies below.

Approach 2: re-analysis of previously published studies of sex differences in regional human brain anatomy

Prior to the availability of large sMRI datasets like the UK Biobank [12], the only way to probe potential sex differences in human neuroanatomy while achieving relatively large sample sizes was to combine the results of previously published studies. However, re-analyses of prior publications suffer from several key limitations as compared to direct analysis of large datasets. Whereas direct data analysis in large samples can provide precise estimates of the effect size for a given sex difference in regional brain anatomy under pre-specified methodological conditions (e.g., methods for segmentation or controlling for brain size), re-analyses of prior publications have to combine findings across diverse studies that vary in sample size and methodology, and are likely to be affected by publication bias towards positive findings [120]. The best way of mitigating sources of heterogeneity and testing for publication bias in review is through meta-analysis, a statistical process for combining data from multiple studies after weighing evidence by sample size to derive summary effect size estimates [6, 7, 121]. However, reliable and interpretable summary effect sizes for neuroanatomical sex differences can only be computed by meta-analysis when there are multiple primary studies that share common approaches to anatomical measurement and statistical control for brain size variation. Unfortunately, the methodological heterogeneity of published sMRI studies on sex differences has precluded meta-analysis across all major study design permutations.

In contrast to the meta-analytic approach, studies that seek to summarize past reports without formal meta-analysis are referred to as “systematic reviews” if they explicitly define literature search terms and study inclusion/exclusion criteria [122] or “reviews” if they do neither of these things. The “meta-synthesis” by Eliot and colleagues [1] is a review that tallies past reports according to whether they found a given anatomical measure to be significantly larger in males, larger in females, or not significantly different between the sexes.

Whether by meta-analysis, systematic review, or review, any attempt to assess sex differences in regional human brain anatomy by combining results of prior research should weight the available evidence by two key study design considerations. First, it is critical to weigh the evidence from past studies by their sample size. This procedure is necessary because of the basic statistical fact that larger sample sizes drawn from a target population provide more accurate and stable estimates of parameters (e.g., the true population mean). By contrast, small

sample sizes may produce spurious results or exaggerate true effect sizes [2]. Second, it is crucial to stratify prior studies by the approach taken for modeling brain size variation. This is because some common techniques for brain size-correction introduce directional biases into the estimation of neuroanatomical sex differences, due to interactions between sex differences in average brain size and the existence of region-specific, non-linear relationships between region volumes and total brain volume [59, 109–111]. For example, the proportion method (i.e., taking a ratio of region size to overall brain size) assumes that all regions take up the same proportion of the brain regardless of brain size (i.e., ‘isometry’). Accordingly, for regions that scale hyper-allometrically with brain size (i.e., take up a larger proportion of larger brains) (e.g., cerebral white matter), proportionalization will artificially decrease region size in females and increase region size in males. Conversely, for regions that scale hypo-allometrically with brain size (i.e., take up a smaller proportion of larger brains) (e.g., amygdala), proportionalization will artificially increase region size in females and decrease region size in males. Additionally, the residual adjustment method (i.e., using the residuals from a regression of region size on brain size) can lead to biased results when brain size is correlated with any of the other predictors (e.g., sex). These biases can create or obscure sex differences in brain anatomy. Such distortions are substantially reduced, or avoided completely, by studies that include brain size as a covariate in multiple regression models. In particular, given that the volume of many brain regions scales non-linearly with brain size, it is recommended to log-transform region and total brain volumes [2]. This transformation is necessary to induce a linear relationship between region volume and overall brain size, and can affect results (see e.g., Williams and colleagues’ [2] for comparisons of linear and non-linear adjustment methods).

Thus, tallying the results of published studies without considering the key methodological factors detailed above would be predicted to suggest highly inconsistent findings regarding sex differences in human brain anatomy, which is precisely what is found in the review by Eliot and colleagues [1].

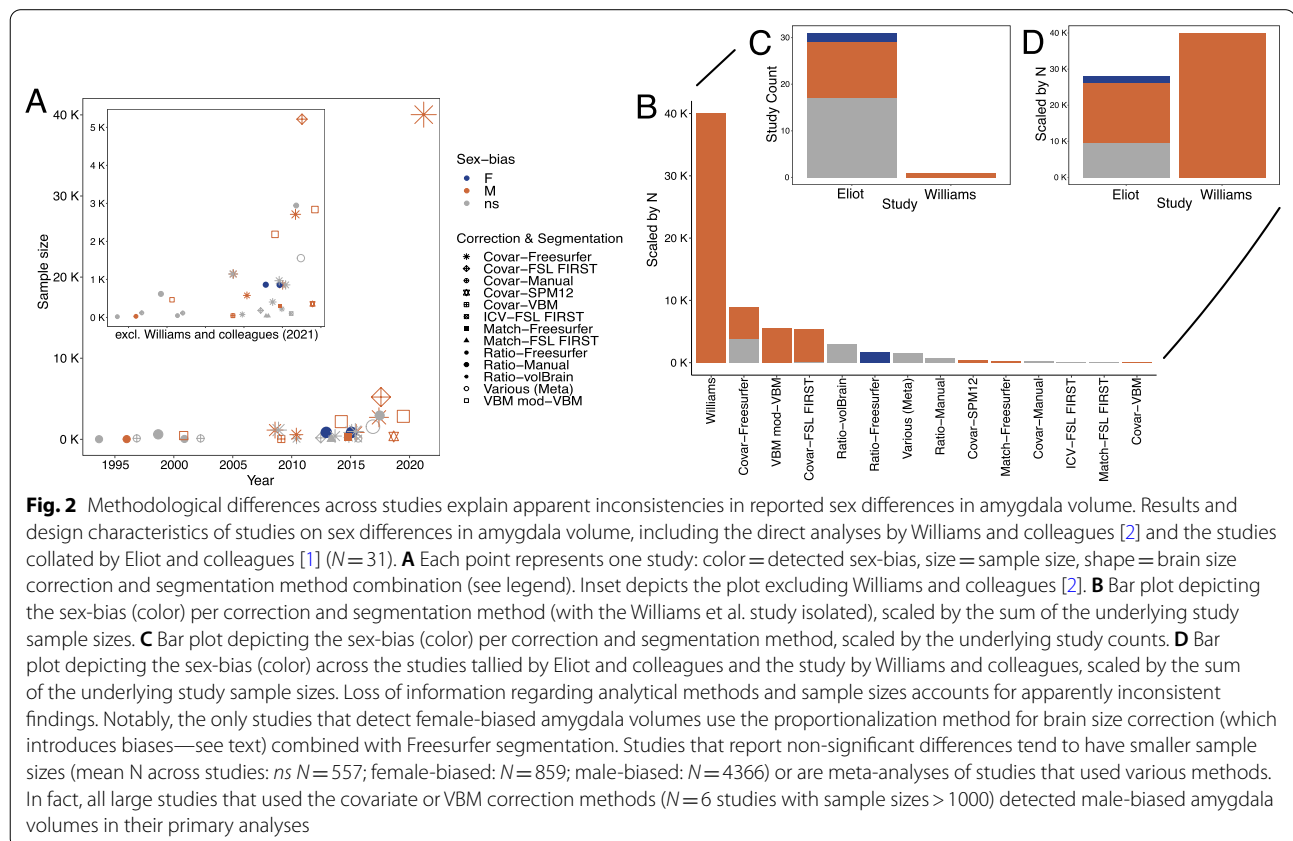
Consideration of scientific study design is crucial when weighing published studies on sex differences in human neuroanatomy

Based on their review of past studies, which collectively included up to ~30k participants, Eliot and colleagues conclude that “once we account for individual differences in brain size, there is almost no difference in the volume of specific cortical or subcortical structures between men and women” [1]. In contrast, based on their direct

analysis of ~40k brain scans, Williams and colleagues conclude that “we find that sex differences in total brain volume are not accounted for by sex differences in height and weight, and that once global brain size is taken into account, there remain numerous regional sex differences in both directions” [3]. The study design considerations reviewed above can help explain this discrepancy.

The overlapping regional sex differences in brain volume across 3 independent samples (Fig. 1, Additional file 1: Table S1 [8, 9],) and the high statistical precision with which Williams and colleagues were able to detect regional anatomical sex differences in the full 40k UKB sample [2] suggest that methodological factors may have led Eliot and colleagues’ literature review to underestimate the consistency of sex differences in regional brain anatomy. This hypothesis is supported by the observations that: i) individual reports collated by Eliot and colleagues varied greatly in their sample sizes and techniques for anatomical analysis; and ii) results of these studies were tallied without adequate adjustment for these study design features. The problem posed by this heterogeneity is well-illustrated by considering Elliot and colleagues’ results tally for the amygdala [1]—a region that consistently demonstrates significant male-biased volume (in the left or both hemispheres) across multiple

recent large-scale studies [2, 8, 9] (Fig. 1, Additional file 1: Table S1), as well as in the study by Williams and colleagues [2]. Figure 2A shows the individual studies of amygdala volume that were tallied by Elliot and colleagues, plotted as a function of date, sample size, and technique for brain size control, with the study by Williams and colleagues included for reference. This plot highlights the gross heterogeneity across these study properties, the sharp recent rise in sample size yielding studies with similar findings, and the order-of-magnitude increase in sample size represented by Williams and colleagues’ recent study. Figure 2C provides a direct visualization of the Eliot and colleagues’ tally of past findings for regional amygdala volume, while Fig. 2B and D re-visualizes this tally with additional information on study sample size and methodology. We also provide a sample size weighted entry for the results of Williams and colleagues [2]. This sample size and methodology-informed re-tallying of the studies reviewed by Eliot and colleagues reveals a strong signal from past literature for mean amygdala volume being larger in males than females after accounting for brain volume using unbiased methods, which is precisely the result reported by direct MRI data analysis in multiple recent large studies including Williams et al. [2, 8, 9].



We provide the above example to illustrate the inherent limitations in using review or meta-analysis across multiple small samples to draw conclusions regarding the existence and reproducibility of sex differences in regional brain anatomy [1]. Many of these issues can be avoided in direct analyses of the large and well-powered neuroimaging datasets (Fig. 1, Additional file 1: Table S1, [2, 8, 9]) that we are now so fortunate to have access to [10]. Careful attention to study design not only helps to resolve apparent inconsistencies throughout past research on sex differences in human neuroanatomy, but also provides an objective, rational framework to help researchers and general readers sensibly weigh any future research on this topic.

Designing studies to understand inconsistent results

Despite the overwhelming consistency across large, well-conducted studies of sex differences in human neuroanatomy (Fig. 1D, Additional file 1: Table S1), there are also some inconsistencies (e.g., Fig. 1E, Additional file 2: Table S2). For example, at the granular voxel-level, we find that 2 or more cohorts examined by the recent large-scale studies of regional sex differences in GMV find opposing statistically significant sex differences for the calcarine sulcus and multiple cerebellar areas (Fig. 1E, Additional file 2: Table S2). These inconsistencies are few but represent opportunities for us to improve our understanding of neuroimaging research on neuroanatomical sex differences in humans. Specifically, targeted research can assess whether a lack of reproducibility in these specific regions reflects particular statistical, technical, or biological factors. Previous studies have highlighted multiple factors that are likely to be critical, including: (i) differences in statistical power due to sample size variation; (ii) the use of different brain size correction methods; (iii) the use of different segmentation approaches; (iv) the use of voxel-wise versus region-wise approaches; (v) region-specific sex differences in age-related size changes; and (vi) subregion-specific sex differences. For instance, Williams and colleagues [2] replicated a previous analysis of a subsample of the UKB ($N \sim 5200$) [123] by applying the log–log regression method (versus the linear covariate adjustment) and an alternative subcortical segmentation approach. Although a majority of the sex effects were replicated, the replication analysis detected sex differences for multiple regions that were not identified in the original study. Similarly, Williams and colleagues [2] found that different cerebellar segmentation approaches produced opposite results (i.e., cerebellar GMV was female-biased in one analysis and male-biased in another), a finding that highlights potential

issues with automated (versus manual) segmentation approaches required by studies with extremely large sample sizes. Variation across segmentation methods may be particularly relevant to structurally complex regions. Furthermore, the use of voxel-wise versus region-wise approaches can lead to discrepancies across studies, since the former may detect sex differences within *part* of a region, but when the sex-biased area does not comprise a large enough portion of the overall region size, no difference is detected by the latter (e.g., the right amygdala is male-biased according to voxel-wise comparisons in [8, 9] but exhibits no bias under region-wise comparisons in [2]). We might also expect to find inconsistent results across studies of different age groups when there are region-specific interactions between sex and age (linear or quadratic) [2, 9]. While such interactions may be relevant to a number of brain areas [2] this issue may be particularly relevant to the hippocampus, a developmentally plastic structure in which adult neurogenesis occurs [124]. Analyses of this region also suggest that sex-by-age interactions vary across different subregions of the hippocampus [2, 9], suggesting that seemingly inconsistent results for functionally and structurally heterogeneous structures (like the hippocampus) may be explained by examining subregion-specific sex differences. Similarly, it would be more appropriate to analyze a region like the hypothalamus on a subregion level, as it is composed of numerous functionally distinct nuclei that are likely to differ in their patterns of sex-biased volume in humans (as in mice [125]). Thus, when large, well-conducted studies generate inconsistent findings regarding sex differences in regional brain anatomy, these inconsistencies should themselves be treated as an object of scientific inquiry to understand the biological or methodological factors that might drive heterogeneous results. Moreover, while seeking to understand inconsistencies, special scrutiny should be applied to results that originate from studies that have smaller samples and use suboptimal methods for brain size control.

Part 4: Framing and interpreting the results

Confusion and ineffective communication around the topic of sex differences in human neuroanatomy is not only shaped by the decisions made en route to generating results, but also by the way these results are discussed and interpreted. Below, we consider important issues in the interpretation of reported sex differences in regional human brain anatomy.

Exploratory versus hypothesis-driven approaches to the study of neuroanatomical sex differences

The observations outlined earlier in this *Commentary* (e.g., sexual selection mechanisms, environmental sex differences, the utility of in vivo sMRI) represent broad hypotheses and assumptions that may guide us towards the exploratory approaches most likely to improve our understanding of human brain sex differences [126]. However, many of these observations do not represent the kinds of specific hypotheses envisioned by most hypothesis-driven accounts of scientific method [126]. The relative value of exploratory versus hypothesis-driven approaches has been debated for decades, and although more targeted analyses (of specific areas) with pre-registered research plans prevent researchers from making post hoc rationalizations for sex differences detected from whole brain screens [17], modern data-intensive scientific disciplines do tend to be characterized by an iterative interplay between hypothesis-driven and exploratory research [126]. This iterative model of scientific practice suggests that any study design can be appropriate if it incorporates approaches and methods that best address gaps in current knowledge, even if the research does not focus on hypothesis testing [126]. While this viewpoint bolsters the utility of exploratory studies of neuroanatomical sex differences, it also highlights the value and complementarity of targeted, hypothesis-driven studies. Researchers implementing either approach should aim to prevent misinterpretations of their findings by explicitly stating the exploratory nature of the work upfront (with appropriate statistical correction for multiple comparisons) and avoiding speculative, functional extrapolations for which there is no evidence (discussed more below).

The complex relationship between sex and gender and the need for intersectional approaches

Studies examining potential sex differences in human brain anatomy should not only be clear regarding their definition of sex (see above), but also recognize the distinct concepts of sex versus gender. Indeed, more studies are explicitly considering participants' sex *and* their gender to more accurately represent individuals on both spectra and recognize the likelihood of separate, compounding, and interacting effects of these variables on disease susceptibility, presentation, and diagnosis [127–129]. Although this approach is still in its infancy, it is likely to represent a critical step towards understanding the entwined social and biological drivers of these conditions. Furthermore, intersectional approaches that consider potential interacting effects of other social identities and categories (e.g., ethnicity) are critical since gender

effects can only be fully understood in the context of these identities.

Identifying sex differences versus classifying individuals by sex

The studies discussed above largely focus on the specific question: Do humans show sex differences in regional brain anatomy by sMRI? This is distinct from the multivariate classification version of this question: Can anatomical variation (captured by e.g., sMRI-derived measures) across multiple brain regions be used to reliably classify the sex of an individual brain? Many studies have demonstrated that multivariate classification algorithms can, in fact, predict sex based on neuroanatomical features with good accuracy [4, 130]. However, this approach can limit our understanding of individual variation within the sexes [5, 131]. Furthermore, these findings should not be interpreted as evidence for large sex differences in brain structure since classification and estimation are very different concepts [132]. Accordingly, identifying specific brain areas that exhibit sex-biased volumes, estimating the magnitude of those average differences, and investigating inter-regional and inter-individual deviations from these averages requires the approach used in many studies discussed here (including those by Williams and colleagues [2, 3]), namely direct analysis of sex differences in regional brain anatomy using large sMRI datasets.

Expectations and interpretations of the magnitude of neuroanatomical sex differences

The increasingly large sample sizes available for studies of human neuroanatomy make it possible to recover statistically significant sex differences in regional brain anatomy that involve small sex differences in mean volume [133]. The small magnitude of most neuroanatomical sex differences is sometimes used to dismiss these differences as unimportant [1], but whether the size of these differences falls below, exceeds, or is in line with prior expectations has rarely been addressed in studies of neuroanatomical sex differences.

Comparative biology can provide some insights that may guide these expectations. Previous work suggests that species with smaller sex differences in behavior and environment also tend to exhibit smaller brain sex differences [23, 24, 134, 135]. For instance, songbird species with male-limited singing show large sex differences in the size of brain song nuclei, while species with male-female duetting tend to exhibit much smaller differences [23]. Similarly, monogamous vole species exhibit small sex differences in home range size, spatial ability, and relative hippocampus size, but these measures are male-biased in polygynous species since only males need to

search for mates [24]. Given that humans exhibit little to no sex differences for many cognitive and psychological traits [136, 137], it follows that we should expect any structural sex differences in the human brain to *also be small* (Fig. 3). Given this expectation of small average sex differences, it also follows that we *should not* expect non-overlapping distributions for neuroanatomical measures between the sexes (i.e., sexual dimorphism). Notably, an observation of overlapping distributions does not reject a hypothesis centered on group-level mean differences. This is consistent across countless non-sex-related study designs created to test whether there is a significant effect of a binary trait (e.g., the Western versus Mediterranean diet) on another, continuous trait (e.g., BMI).

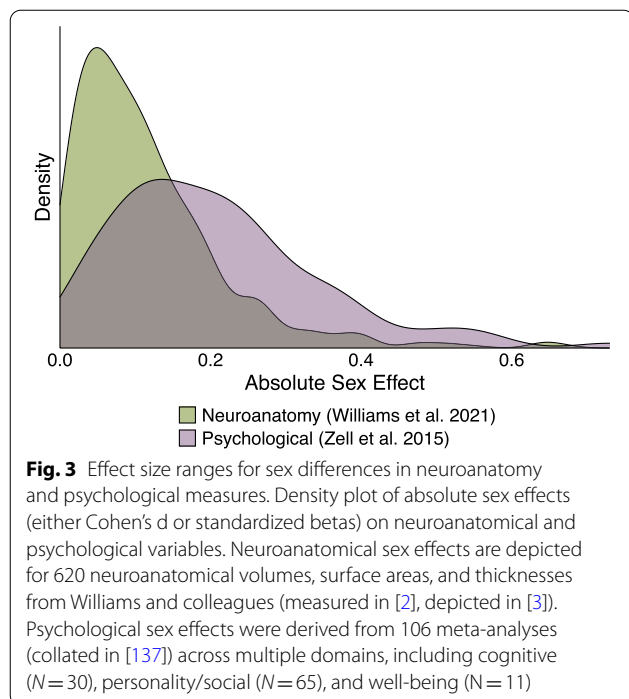
Nevertheless, neuroanatomical sex differences should not simply be dismissed because they are “small”. In fact, previous work suggests that integrating multiple individual measures (each with small effect sizes) into a multivariate measure can produce a much larger effect size that implies statistical separation of males and females (e.g., sex differences in some personality traits and aggression) [138, 139]. Furthermore, there is no a priori reason to assume that large, widespread structural differences are required to produce the observed sex differences in disease prevalence and presentation (discussed above). The magnitude of biological and/or functional implications of empirically small anatomical differences cannot be assumed based on the latter. For example, duplication of the Y chromosome represents

an empirically small amount of excess genomic material (chrY = ~2% of the genome), but it can result in large neuroanatomical and behavioral alterations (e.g., ~18% increase in ASD prevalence; ~67% increase in special educational needs; both vs. national prevalence rates in England) [140]. Similarly, small volumetric differences may reflect empirically larger differences in finer aspects of neurobiology (e.g., cell counts) [101] that may be relevant to disease susceptibility. Accordingly, *small average* sex differences in brain structure may potentially help to illuminate the mechanisms underlying differences in disease, either directly or by revealing regional sex differences in other levels of biological organization that are proxied by anatomy.

Causes and effects of neuroanatomical sex differences

Studies that describe a sex difference in average region volume do not provide any information regarding the causal bases for that difference. In particular, these studies cannot illuminate which microstructural features underpin the observed anatomical differences, which proximal genetic and/or environmental factors may establish these differences, or which (if any) distal evolutionary factors were at play throughout early human evolution. Rather, as detailed above, the value of sMRI studies is that they help to prioritize brain systems for closer investigation using other selected research approaches.

Previous work suggests that structural sex differences may reflect underlying differences that cannot be directly inferred from volume and that may even exhibit the opposite-sex bias [101]. Accordingly, sMRI-directed postmortem techniques would be required to study the microstructural bases for in vivo differences in regional brain volume in humans. In contrast, experimental animal studies would be needed to probe potential mechanistic drivers of sex-biased brain development [42, 141], with the attendant caveats that arise when using certain model systems to study human phenotypes [142, 143]. Such research would also need to explicitly grapple with the multi-level nature of brain organization, the possibility that sex differences may show a complex relationship between different levels of biological analysis, and the fact that many detected sex differences are likely to represent compensatory mechanisms that facilitate equal functionality across individuals [80]. Finally, an understanding of the evolutionary drivers of neuroanatomical sex differences in humans may never be fully within our grasp, but there is potential for comparative studies to provide some insights (for non-primate examples, see: [134, 144]). This approach will require much larger sex-specific neuroanatomical (e.g., sMRI-derived) and behavioral



data sets (both within and across primate species) than are currently available.

Furthermore, the act of simply describing a sex difference in regional brain anatomy is entirely distinct from understanding the functional implications (if any) of such a difference. This not only reflects a lack of understanding of the neurobiological drivers of regional volumetric differences (see above), but the wider complexity of structural and functional relationships between brain regions, as well as how such relationships relate to inter-individual differences in behavior or cognition. While published reports of links between neuroimaging sex differences and sex differences in cognitive or behavioral traits do exist (reviewed in [145]), the neuroimaging community is currently grappling with difficult general questions regarding the degree to which *in vivo* structural and functional features from MRI can predict inter-individual variation in human behavior or cognition [146]. In studies of neuroanatomical sex differences, these questions are compounded further since there are many diverse configurations that could theoretically underlie sex differences in brain–behavior relationships. For example, given a hypothetical dimension of behavior that correlates with neuroanatomical variation within a network of brain regions, males and females may differ in: (i) the set of brain regions that constitute this network; (ii) the strength of brain–behavior correlations for individual regions; (iii) the relationship between behavior and anatomy across multiple regions; and (iv) the specific anatomical feature (e.g., cortical volume vs. cortical thickness) that correlates with behavior. Furthermore, for any single region showing anatomy–behavior correlations (e.g., for region volume relative to brain size), there are many possible combinations of: (i) the presence and direction of sex difference in mean volume; (ii) the presence and direction of sex differences in behavior; and (iii) the valence of brain–behavior relationships (e.g., “greater volume, greater proficiency” vs. “greater volume, less proficiency”). These possibilities alone represent a vast and complex analytic search space, which increases exponentially if considering multiple brain regions and/or behaviors, and faster still if considering developmental changes rather than cross-sectional variations in anatomy. We outline this theoretical complexity to underline how challenging it would be to test the functional relevance of any observed sex differences in brain anatomy. In the absence of specially tailored study designs to deal with this complexity, there is no empirical basis for arbitrating as to whether a given sex difference in regional brain anatomy may or may not have functional relevance. Prior to initiating research in this field, we suggest that researchers consider whether their personal ethics align

with the potential societal implications of this type of work [136, 147].

Careful presentation and framing of findings on sex differences in human brain anatomy—especially with regard to any potential functional implications—may help to limit wider misinterpretations that erroneously deflate or inflate their significance. This caution is especially important given historical precedents for biological data having been misused to justify or explain social inequities [13–15], and the modern risk for discussions of biological sex differences to be unhelpfully simplified, charged, amplified and binarized by social media [148]. Researchers can take proactive steps to decrease the likelihood of such adverse outcomes by explicitly outlining the limits of findings on first publication, including what they do and do not establish, and which hypothesis should be prioritized in follow-up research to replicate and contextualize findings. For example, multiple studies have demonstrated that males exhibit more variability in brain size and structure [116, 117, 123, 149]. In some cases, these results have been misused to justify current gender disparities in STEM fields (due to a higher frequency of males at the “top-end” of neuroanatomical distributions) [150]. However, in other cases, researchers have proactively preempted such biased interpretations by thoughtfully considering the full space of theoretical possibilities and engaging with this space through scientific data-based inquiry. For example: (i) while O’Dea and colleagues [150] confirm that girls exhibit higher average, but less variable grades, they note that by the time girls graduates, they are just as likely as boys to have earned high enough grades to pursue a career in STEM; and (ii) while Wierenga and colleagues [116] confirm greater neuroanatomical variability among men, they note that extreme brain structure (in either direction) may be costly due to energy-induced tradeoffs between volume and conduction time [151].

Following these examples, we argue that any evidence for structural sex differences in the human brain should be discussed from simultaneously scientific and anti-sexist viewpoints. Presenting results on sex-biased brain anatomy from an anti-sexist viewpoint not only requires a mindful avoidance of functional inferences that lack evidence, but also an active effort to highlight the lack of such evidence and carefully define the limits of one’s results. Exercising such caution in functional inferences is fully compatible with an impartial consideration of available data on structural differences.

Conclusions

As detailed above, the best available evidence to date—generated from a direct analysis of tens of thousands of individuals [2, 3] and direct comparisons of results

across independent large cohorts [8, 9]—suggests that the human brain shows highly reproducible sex differences in regional brain anatomy above and beyond sex differences in overall brain size. These are differences in the mean value of largely overlapping distributions, and they show small-to-moderate effect sizes; however, we do not understand their microstructural basis, the causal factors shaping them, or if they facilitate sex differences or equivalences in behavior and cognition. There is also a pressing need to better understand the developmental unfolding of sex differences in human brain anatomy over time. This work will benefit from the recent expansion of large longitudinal neuroimaging datasets of human development, as well as recent advances in the methods available to model potentially dynamic sex differences in these datasets [152].

The existence of many open questions, such as those outlined above, does not undo the existence of anatomical sex differences, but instead identifies important topics for future research, along one of the many paths that will hopefully lead us to a richer and more equitable understanding of human diversity [104]. In pursuing these paths, it will be crucial to actively safeguard impartiality and objectivity in our work. Such protections are critical to all domains of scientific inquiry, but they are especially so for studies of sex differences in human neuroanatomy. This call for special care reflects the existence of pervasive gender-biases throughout society, in addition to an undeniable history of cherry-picked biological data being misappropriated and misrepresented to support sexist points of view [13–15]. Modern science is not isolated from this history or from society. In particular, the act of communicating research findings to the public necessarily engages with societal biases. Indeed, scientific organizations and publishers have enacted mechanisms to safeguard against such biases coloring the conduct or presentation of research [153, 154]. Therefore, adopting an "anti-sexist" viewpoint (i.e., one that proactively counters sex/gender-based prejudice, discrimination, or expectations guided by stereotyping) helps to insulate the scientific method from potential biases that might otherwise lead towards denials or over-interpretations of research on sex differences in human brain anatomy. Rather than avoiding, dismissing, or over-interpreting findings of brain sex differences, more accurate description could reduce the misrepresentation and misuse of such research, both within the scientific community and throughout society as a whole.

Abbreviations

ADHD: Attention deficit disorder; ASD: Autism spectrum disorder; HCP: Human Connectome Project; OCD: Obsessive compulsive disorder; SHIP: Study of health in Pomerania; sMRI: Structural magnetic resonance imaging; STEM:

Science, technology, engineering, math; UKB: UK Biobank; XCI: X chromosome inactivation.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13293-022-00448-w>.

Additional file 1: Table S1. Consistencies in the spatial patterning of neuroanatomical sex differences. This table includes regions showing significant (Family Wise Error < 0.05, corrected) GMV sex differences with the same direction in all three contrasts (Fig. 1D, shown in yellow and cyan). These regions are listed in descending order of cluster size. X, Y and Z coordinates indicate the center-of-mass of each region in the MNI space. Anatomical labels were annotated based on the Automated Anatomical Labeling (AAL) atlas [155]. Regions that are below 20 voxels (that is, 0.0675 cm³) are excluded in this table for simplification. L = left; R = right; partition 1/2 = two separate partitions in one region.

Additional file 2: Table S2. Inconsistencies in the spatial patterning of neuroanatomical sex differences. This table includes regions showing significant (Family Wise Error < 0.05, corrected) GMV sex differences with the opposite direction in any two of three contrasts (Fig. 1E). These regions are listed in descending order of cluster size. X, Y and Z coordinates indicate the center-of-mass of each region in the MNI space. Anatomical labels were annotated based on the Automated Anatomical Labeling (AAL) atlas [155]. Regions that are below 20 voxels (that is, 0.0675 cm³) are excluded in this table for simplification. L = left; R = right; partition 1/2 = two separate partitions in one region.

Acknowledgements

We thank Dr. Martin Lotze for providing access to data for Fig. 1, and Amber Trujillo, Marialanna Lee and Matthew McCoy for their feedback on this manuscript.

Author contributions

ARD and AR conceptualized the *Commentary*. SL and ARD generated the Figures. All authors reviewed and revised the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by the intramural research program of the National Institute of Mental Health (project funding: 1ZIAMH002949-03; clinical trials identifier: NCT00001246; protocol: 89-M-0006) and the National Institute of Child Health and Disease (R01HD100298).

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interest.

Received: 22 February 2022 Accepted: 23 May 2022

Published online: 26 July 2022

References

1. Eliot L, Ahmed A, Khan H, Patel J. Dump the "dimorphism": comprehensive synthesis of human brain studies reveals few male-female differences beyond size. *Neurosci Biobehav Rev.* 2021;125:667–97.

2. Williams CM, Peyre H, Toro R, Ramus F. Neuroanatomical norms in the UK Biobank: the impact of allometric scaling, sex, and age. *Hum Brain Mapp.* 2021;42:4623–42.
3. Williams CM, Peyre H, Toro R, Ramus F. Sex differences in the brain are not reduced to differences in body size. *Neurosci Biobehav Rev.* 2021;130:509–11.
4. Del Giudice M. Binary thinking about the sex binary: a comment on Joel (2021). *Neurosci Biobehav Rev Pergamon.* 2021. <https://doi.org/10.1016/j.neubiorev.2021.04.020>.
5. Joel D. Beyond the binary: rethinking sex and the brain. *Neurosci Biobehav Rev.* 2021. <https://doi.org/10.1016/j.neubiorev.2020.11.018>.
6. Tan A, Ma W, Vira A, Marwha D, Eliot L. The human hippocampus is not sexually-dimorphic: meta-analysis of structural MRI volumes. *Neuroimage.* 2016;124:350–66.
7. Marwha D, Halari M, Eliot L. Meta-analysis reveals a lack of sexual dimorphism in human amygdala volume. *Neuroimage.* 2017;147:282–94.
8. Liu S, Seidlitz J, Blumenthal JD, Clasen LS, Raznahan A. Integrative structural, functional, and transcriptomic analyses of sex-biased brain organization in humans. *Proc Natl Acad Sci.* 2020. <https://doi.org/10.1073/pnas.1919091117>.
9. Lotze M, Domin M, Gerlach FH, Gaser C, Lueders E, Schmidt CO, et al. Novel findings from 2,838 adult brains on sex differences in gray matter brain volume. *Sci Rep.* 2019;9:1671.
10. Madan CR. Scan once, analyse many: using large open-access neuroimaging datasets to understand the brain. *Neuroinformatics.* 2021. <https://doi.org/10.1007/s12021-021-09519-6>.
11. Lerch JP, van der Kouwe AJW, Raznahan A, Paus T, Johansen-Berg H, Miller KL, et al. Studying neuroanatomy using MRI. *Nat Neurosci.* 2017;20:314–26.
12. Miller KL, Alfaro-Almagro F, Bangarter NK, Thomas DL, Yacoub E, Xu J, et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci.* 2016;19:1523–36.
13. Russett CE. *Sexual Science.* Harvard University Press; 2009.
14. Gould SJ, Gold SJ. *The alexander Agassiz professor of zoology Stephen Jay Gould. The Mismeasure of Man.* New York: W. W. Norton & Company; 1996.
15. Saini A. *Inferior: how science got women wrong and the new research that's rewriting the story.* Boston: Beacon Press; 2017.
16. Clayton JA. Applying the new SABV (sex as a biological variable) policy to research and clinical care. *Physiol Behav.* 2018;187:2–5.
17. Rippon G, Eliot L, Genon S, Joel D. How hype and hyperbole distort the neuroscience of sex differences. *PLoS Biol.* 2021;19: e3001253.
18. Tinbergen N. On aims and methods of Ethology. *Z Tierpsychol.* 1963;20:410–33.
19. Mayr E. Cause and effect in biology: kinds of causes, predictability, and teleology are viewed by a practicing biologist. *Science.* 1961;134:1501–6.
20. Laland KN, Sterelny K, Odling-Smee J, Hoppitt W, Uller T. Cause and effect in biology revisited: is mayr's proximate-ultimate dichotomy still useful? *Science.* 2011;334:1512–6.
21. Scott-Phillips TC, Dickins TE, West SA. Evolutionary theory and the ultimate-proximate distinction in the human behavioral sciences. *Perspect Psychol Sci.* 2011;6:38–47.
22. Darwin CR. *The descent of man and selection in relation to sex.* London: Murray; 1871.
23. MacDougall-Shackleton SA, Ball GF. Comparative studies of sex differences in the song-control system of songbirds. *Trends Neurosci.* 1999;22:432–6.
24. Sherry DF, Jacobs LF, Gaulin SJ. Spatial memory and adaptive specialization of the hippocampus. *Trends Neurosci.* 1992;15:298–303.
25. Sherwood CC, Subiaul F, Zawadzki TW. A natural history of the human mind: tracing evolutionary changes in brain and cognition. *J Anat.* 2008;212:426–54.
26. Plavcan JM. Sexual size dimorphism, canine dimorphism, and male-male competition in primates. *Hum Nat.* 2012;23:45–67.
27. Dixon BJ, Brooks RC. The role of facial hair in women's perceptions of men's attractiveness, health, masculinity and parenting abilities. *Evol Hum Behav.* 2013;34:236–41.
28. Puts D. Human sexual selection. *Curr Opin Psychol.* 2016;7:28–32.
29. Lindenfors P, Tullberg BS. Evolutionary aspects of aggression: the importance of sexual selection. *Adv Genet Elsevier.* 2011;75:7–22.
30. Levine SC, Foley A, Lourenco S, Ehrlich S, Ratliff K. Sex differences in spatial cognition: advancing the conversation. *Wiley Interdiscip Rev Cogn Sci.* 2016;7:127–55.
31. Masclans A, Hamon C, Jeunesse C, Bickle P. A sexual division of labour at the start of agriculture? A multi-proxy comparison through grave good stone tool technological and use-wear analysis. *PLoS ONE.* 2021;16: e0249130.
32. Silverman I, Choi J, Peters M. The hunter-gatherer theory of sex differences in spatial abilities: data from 40 countries. *Arch Sex Behav Springer.* 2007;36:261–8.
33. Silverman I, Eals M. *Sex differences in spatial abilities: evolutionary theory and data.* Oxford: Oxford University Press; 1992.
34. Sjöberg EA, Cole GG. Sex differences on the Go/No-Go test of inhibition. *Arch Sex Behav Springer.* 2018;47:537–42.
35. McCarthy MM, Arnold AP, Ball GF, Blaustein JD, De Vries GJ. Sex differences in the brain: the not so inconvenient truth. *J Neurosci.* 2012;32:2241–7.
36. Hiramatsu R, Matoba S, Kanai-Azuma M, Tsunekawa N, Katoh-Fukui Y, Kurohmaru M, et al. A critical time window of Sry action in gonadal sex determination in mice. *Development.* 2009;136:129–38.
37. Goodfellow PN, Lovell-Badge R. SRY and sex determination in mammals. *Annu Rev Genet.* 1993;27:71–92.
38. Bae YJ, Zeidler R, Baber R, Vogel M, Wirkner K, Loeffler M, et al. Reference intervals of nine steroid hormones over the life-span analyzed by LC-MS/MS: Effect of age, gender, puberty, and oral contraceptives. *J Steroid Biochem Mol Biol.* 2019;193: 105409.
39. Phoenix CH, Goy RW, Gerall AA, Young WC. Organizing action of prenatally administered testosterone propionate on the tissues mediating mating behavior in the female guinea pig. *Endocrinology.* 1959;65:369–82.
40. Morris JA, Jordan CL, Breedlove SM. Sexual differentiation of the vertebrate nervous system. *Nat Neurosci.* 2004;7:1034–9.
41. Wallen K. Hormonal influences on sexually differentiated behavior in nonhuman primates. *Front Neuroendocrinol.* 2005;26:7–26.
42. Corre C, Friedel M, Vousden DA, Metcalf A, Spring S, Qiu LR, et al. Separate effects of sex hormones and sex chromosomes on brain structure and function revealed by high-resolution magnetic resonance imaging and spatial navigation assessment of the Four Core Genotype mouse model. *Brain Struct Funct Springer.* 2016;221:997–1016.
43. Vousden DA, Corre C, Spring S, Qiu LR, Metcalf A, Cox E, et al. Impact of X/Y genes and sex hormones on mouse neuroanatomy. *Neuroimage.* 2018;173:551–63.
44. Liao Z, Patel Y, Khairullah A, Parker N, Paus T. Pubertal testosterone and the structure of the cerebral cortex in young men. *Cereb Cortex.* 2021;31:2812–21.
45. Peper JS, Hulshoff Pol HE, Crone EA, van Honk J. Sex steroids and brain structure in pubertal boys and girls: a mini-review of neuroimaging studies. *Neuroscience.* 2011;191:28–37.
46. Dubol M, Epperson CN, Sacher J, Pletzer B, Derntl B, Lanzenberger R, et al. Neuroimaging the menstrual cycle: a multimodal systematic review. *Front Neuroendocrinol.* 2021;60: 100878.
47. Merke DP, Fields JD, Keil MF, Vaituzis AC, Chrousos GP, Giedd JN. Children with classic congenital adrenal hyperplasia have decreased amygdala volume: potential prenatal and postnatal hormonal effects. *J Clin Endocrinol Metab.* 2003;88:1760–5.
48. Kranz GS, Zhang BBB, Handschuh P, Ritter V, Lanzenberger R. Gender-affirming hormone treatment—A unique approach to study the effects of sex hormones on brain structure and function. *Cortex Elsevier.* 2020;129:68–79.
49. Arnold AP. The end of gonad-centric sex determination in mammals. *Trends Genet.* 2012;28:55–61.
50. Raznahan A, Disteche CM. X-chromosome regulation and sex differences in brain anatomy. *Neurosci Biobehav Rev.* 2021;120:28–47.
51. Oliva M, Muñoz-Agüirre M, Kim-Hellmuth S, Wucher V, Gewirtz ADH, Cotter DJ, et al. The impact of sex on gene expression across human tissues. *Science.* 2020. <https://doi.org/10.1126/science.aba3066>.
52. Berletch JB, Yang F, Xu J, Carrel L, Disteche CM. Genes that escape from X inactivation. *Hum Genet.* 2011;130:237–45.
53. McCarthy MM, Nugent BM, Lenz KM. Neuroimmunology and neuroepigenetics in the establishment of sex differences in the brain. *Nat Rev Neurosci.* 2017;18:471–84.

54. Raznahan A, Parikshak NN, Chandran V, Blumenthal JD, Clasen LS, Alexander-Bloch AF, et al. Sex-chromosome dosage effects on gene expression in humans. *Proc Natl Acad Sci U S A*. 2018;115:7398–403.
55. Bermejo-Alvarez P, Rizos D, Rath D, Lonergan P, Gutierrez-Adan A. Sex determines the expression level of one third of the actively expressed genes in bovine blastocysts. *Proc Natl Acad Sci U S A*. 2010;107:3394–9.
56. Dewing P, Shi T, Horvath S, Vilain E. Sexually dimorphic gene expression in mouse brain precedes gonadal differentiation. *Mol Brain Res*. 2003;118:82–90.
57. De Vries GJ, Rissman EF, Simerly RB, Yang LY, Scordalakes EM, Auger CJ, et al. A model system for study of sex chromosome effects on sexually dimorphic neural and behavioral traits. *J Neurosci*. 2002;22:9005–14.
58. Raznahan A, Lee NR, Greenstein D, Wallace GL, Blumenthal JD, Clasen LS, et al. Globally divergent but locally convergent X- and Y-chromosome influences on cortical development. *Cereb Cortex*. 2016;26:70–9.
59. Mankiw C, Park MTM, Reardon PK, Fish AM, Clasen LS, Greenstein D, et al. Allometric analysis detects brain size-independent effects of sex and sex chromosome complement on human cerebellar organization. *J Neurosci*. 2017;37:5221–31.
60. Fish AM, Nadig A, Seidlitz J, Reardon PK, Mankiw C, McDermott CL, et al. Sex-biased trajectories of amygdalo-hippocampal morphology change over human development. *Neuroimage*. 2020;204: 116122.
61. Warling A, Yavi M, Clasen LS, Blumenthal JD, Lalonde FM, Raznahan A, et al. Sex chromosome dosage effects on white matter structure in the human brain. *Cereb Cortex*. 2021;31:5339–53.
62. Jensen MD. Adipose tissue and fatty acid metabolism in humans. *J R Soc Med*. 2002;95(Suppl 42):3–7.
63. Roy AK, Chatterjee B. Sexual dimorphism in the liver. *Annu Rev Physiol*. 1983;45:37–50.
64. Mirandola L, Wade R, Verma R, Pena C, Hosiriluck N, Figueroa JA, et al. Sex-driven differences in immunological responses: challenges and opportunities for the immunotherapies of the third millennium. *Int Rev Immunol*. 2015;34:134–42.
65. Liu J, Sharma N, Zheng W, Ji H, Tam H, Wu X, et al. Sex differences in vasopressin V2 receptor expression and vasopressin-induced antidiuresis. *Am J Physiol Renal Physiol*. 2011;300:F433–40.
66. de Vries GJ, Forger NG. Sex differences in the brain: a whole body perspective. *Biol Sex Differ*. 2015;6:15.
67. Cárdenas EF, Kujawa A, Humphreys KL. Neurobiological changes during the peripartum period: implications for health and behavior. *Soc Cogn Affect Neurosci*. 2020;15:1097–110.
68. Hoekzema E, Barba-Müller E, Pozzobon C, Picado M, Lucco F, García-García D, et al. Pregnancy leads to long-lasting changes in human brain structure. *Nat Neurosci*. 2017;20:287–96.
69. Horrell ND, Acosta MC, Saltzman W. Plasticity of the paternal brain: Effects of fatherhood on neural structure and function. *Dev Psychobiol*. 2021;63:1499–520.
70. Schwab K. The global gender gap report 2020. Cologne: World Economic Forum; 2019.
71. UNESCO. Global education monitoring report 2019: gender report: Building bridges for gender equality. Paris: UNESCO Publishing; 2019. p. 2019.
72. Huang J, Krivkovich A, Starikova I, Yee L, Zanoschi D. Women in the Workplace 2019. McKinsey & Company [Internet]. 2019; <https://www.agec.org.au/wp-content/uploads/2019/10/McKinsey-Women-in-the-workplace-2019.pdf>.
73. Quinn MM, Smith PM. Corrigendum: gender, work, and health. *Ann Work Expo Health*. 2018;62:516.
74. Glassman B. An analysis of the gender poverty gap using the American Community Survey [Internet]. 2020. <https://www.census.gov/content/dam/Census/library/working-papers/2020/demo/sehdsd-wp2020-12.pdf>. Accessed 2 Feb 2022.
75. Phillips SP. Defining and measuring gender: a social determinant of health whose time has come. *Int J Equity Health*. 2005;4:11.
76. Hanson RF, Borntreger C, Self-Brown S, Kilpatrick DG, Saunders BE, Resnick HS, et al. Relations among gender, violence exposure, and mental health: the national survey of adolescents. *Am J Orthopsychiatry*. 2008;78:313–21.
77. Caspi A, Moffitt TE. Gene-environment interactions in psychiatry: Joining forces with neuroscience. *Nat Rev Neurosci*. 2006;7:583–90.
78. Eliot L. Brain development and physical aggression: how a small gender difference grows into a violence problem. *Curr Anthropol*. 2021;62:566–78.
79. Burgaleta M, Head K, Álvarez-Linera J, Martínez K, Escorial S, Haier R, et al. Sex differences in brain volume are related to specific skills, not to general intelligence. *Intelligence*. 2012;40:60–8.
80. De Vries GJ. Minireview: Sex differences in adult and developing brains: compensation, compensation, compensation. *Endocrinology*. 2004;145:1063–8.
81. Rutter M, Caspi A, Moffitt TE. Using sex differences in psychopathology to study causal mechanisms: unifying issues and research strategies. [Review] [219 refs]. *J Child Psychol Psychiatry*. 2003. <https://doi.org/10.1111/1469-7610.00194>
82. Ullah MF, Ahmad A, Bhat SH, Abu-Duhier FM, Barreto GE, Ashraf GM. Impact of sex differences and gender specificity on behavioral characteristics and pathophysiology of neurodegenerative disorders. *Neurosci Biobehav Rev*. 2019;102:95–105.
83. Merikangas AK, Almasy L. Using the tools of genetic epidemiology to understand sex differences in neuropsychiatric disorders. *Genes Brain Behav*. 2020;19: e12660.
84. Halladay AK, Bishop S, Constantino JN, Daniels AM, Koenig K, Palmer K, et al. Sex and gender differences in autism spectrum disorder: summarizing evidence gaps and identifying emerging areas of priority. *Mol Autism*. 2015;6:36.
85. Ramtekkar UP, Reiersen AM, Todorov AA, Todd RD. Sex and age differences in attention-deficit/hyperactivity disorder symptoms and diagnoses: implications for DSM-V and ICD-11. *J Am Acad Child Adolesc Psychiatry*. 2010;49:217–28.e1-3.
86. Mathes BM, Morabito DM, Schmidt NB. Epidemiological and Clinical Gender Differences in OCD. *Curr Psychiatry Rep*. 2019;21:36.
87. Maughan B, Collishaw S, Stringaris A. Depression in childhood and adolescence. *J Can Acad Child Adolesc Psychiatry*. 2013;22:35–40.
88. Copeland WE, Angold A, Shanahan L, Costello EJ. Longitudinal patterns of anxiety from childhood to adulthood: the Great Smoky Mountains Study. *J Am Acad Child Adolesc Psychiatry*. 2014;53:21–33.
89. Smink FRE, van Hoeken D, Hoek HW. Epidemiology of eating disorders: incidence, prevalence and mortality rates. *Curr Psychiatry Rep*. 2012;14:406–14.
90. Mazure CM, Swendsen J. Sex differences in Alzheimer's disease and other dementias. *Lancet Neurol*. 2016;15:451–2.
91. Gogos A, Ney LJ, Seymour N, Van Rheenen TE, Felmingham KL. Sex differences in schizophrenia, bipolar disorder, and post-traumatic stress disorder: are gonadal hormones the link? *Br J Pharmacol*. 2019;176:4119–35.
92. Möller-Leimkühler AM. Barriers to help-seeking by men: a review of sociocultural and clinical literature with particular reference to depression. *J Affect Disord*. 2002;71:1–9.
93. Hull L, Petrides KV, Mandy W. The female autism phenotype and camouflaging: a narrative review. *Rev J Autism Dev Disord*. 2020;7:1–12.
94. Kissel LT, Werling DM. Neural transcriptomic analysis of sex differences in autism spectrum disorder: current insights and future directions. *Biol Psychiatry*. 2022. <https://doi.org/10.1016/j.biopsych.2020.11.023>.
95. Green T, Flash S, Reiss AL. Sex differences in psychiatric disorders: what we can learn from sex chromosome aneuploidies. *Neuropsychopharmacology*. 2019;44:9–21.
96. Shulman LM. Gender differences in Parkinson's disease. *Gend Med*. 2007;4:8–18.
97. Abel KM, Drake R, Goldstein JM. Sex differences in schizophrenia. *Int Rev Psychiatry*. 2010;22:417–28.
98. Ferretti MT, Iulita MF, Cavado E, Chiesi PA, Schumacher Dimech A, Santuocione Chadha A, et al. Sex differences in Alzheimer disease—the gateway to precision medicine. *Nat Rev Neurol*. 2018;14:457–69.
99. Raznahan A, Lue Y, Probst F, Greenstein D, Giedd J, Wang C, et al. Triangulating the sexually dimorphic brain through high-resolution neuroimaging of murine sex chromosome aneuploidies. *Brain Struct Funct*. 2015;220:3581–93.
100. Qiu LR, Fernandes DJ, Szulc-Lerch KU, Dazai J, Nieman BJ, Turnbull DH, et al. Mouse MRI shows brain areas relatively larger in males emerge before those larger in females. *Nat Commun*. 2018;9:2615.
101. Kim Y, Yang GR, Pradhan K, Venkataraju KU, Bota M, García Del Molino LC, et al. Brain-wide maps reveal stereotyped cell-type-based cortical

- architecture and subcortical sexual dimorphism. *Cell*. 2017;171:456–469.e22.
102. Wilson MA, Buetow KH. Novel mechanisms of cancer emerge when accounting for sex as a biological variable. *Cancer Res*. 2020;80:27–9.
 103. Astorino CM. Beyond dimorphism: Sexual polymorphism and research bias in biological anthropology. *Am Anthropol*. 2019;121:489–90.
 104. Cahill L. Why sex matters for neuroscience. *Nat Rev Neurosci*. 2006;7:477–84.
 105. Raznahan A, Lee Y, Stidd R, Long R, Greenstein D, Clasen L, et al. Longitudinally mapping the influence of sex and androgen signaling on the dynamics of human cortical maturation in adolescence. *Proc Natl Acad Sci U S A*. 2010;107:16988–93.
 106. DeCasien AR, Higham JP. Relative cerebellum size is not sexually dimorphic across primates. *Brain Behav Evol*. 2020;95:93–101.
 107. Dekaban AS. Changes in brain weights during the span of human life: relation of brain weights to body heights and body weights. *Ann Neurol*. 1978;4:345–56.
 108. Finlay BL, Darlington R. Linked regularities in the development and evolution of mammalian brains. *Science*. 1995;268:1578–84.
 109. Nadig A, Reardon PK, Seidlitz J, McDermott CL, Blumenthal JD, Clasen LS, et al. Carriage of supernumerary sex chromosomes decreases the volume and alters the shape of limbic structures. *BioRxiv*. 2018. <https://doi.org/10.1523/ENEURO.0265-18.2018>.
 110. Reardon PK, Clasen L, Giedd JN, Blumenthal J, Lerch JP, Chakravarty MM, et al. An allometric analysis of sex and sex chromosome dosage effects on subcortical anatomy in humans. *J Neurosci*. 2016;36:2438–48.
 111. Reardon PK, Seidlitz J, Vandekar S, Liu S, Patel R, Park MTM, et al. Normative brain size variation and brain shape diversity in humans. *Science*. 2018;360:1222–7.
 112. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods*. 2011;8:665–70.
 113. Karnath H-O, Sperber C, Rorden C. Mapping human brain lesions and their functional consequences. *Neuroimage*. 2018;165:180–9.
 114. LeWinn KZ, Sheridan MA, Keyes KM, Hamilton A, McLaughlin KA. Sample composition alters associations between age and brain structure. *Nat Commun*. 2017;8:874.
 115. Bhagwat N, Barry A, Dickie EW, Brown ST, Devenyi GA, Hatano K, et al. Understanding the impact of preprocessing pipelines on neuroimaging cortical surface analyses. *Gigascience*. 2021. <https://doi.org/10.1093/gigascience/giaa155>.
 116. Wierenga LM, Doucet GE, Dima D, Agartz I, Aghajani M, Akudjedu TN, et al. Greater male than female variability in regional brain structure across the lifespan. *Hum Brain Mapp*. 2020. <https://doi.org/10.1002/hbm.25204>.
 117. Forde NJ, Jayachandra J, Joseph M, Jacobs GR, Dickie E, Satterthwaite TD, et al. Sex differences in variability of brain structure across the lifespan. *Cereb Cortex*. 2020;30:5420–30.
 118. Goldstein JM. Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. *Cereb Cortex*. 2001;11:490–7.
 119. Caviness VS Jr, Meyer J, Makris N, Kennedy DN. MRI-based topographic parcellation of human neocortex: an anatomically specified method with estimate of reliability. *J Cogn Neurosci*. 1996;8:566–87.
 120. Joobar R, Schmitz N, Annable L, Boksa P. Publication bias: what are the challenges and can they be overcome? *J Psychiatry Neurosci*. 2012;37:149–52.
 121. Ruigrok ANV, Salimi-Khorshidi G, Lai M-C, Baron-Cohen S, Lombardo MV, Tait RJ, et al. A meta-analysis of sex differences in human brain structure. *Neurosci Biobehav Rev*. 2014;39:34–50.
 122. Gopalakrishnan S, Ganeshkumar P. Systematic reviews and meta-analysis: understanding the best evidence in primary healthcare. *J Family Med Prim Care*. 2013;2:9–14.
 123. Ritchie SJ, Cox SR, Shen X, Lombardo MV, Reus LM, Alloza C, et al. Sex differences in the adult human brain: evidence from 5216 UK biobank participants. *Cereb Cortex*. 2018;28:2959–75.
 124. Leuner B, Gould E. Structural plasticity and hippocampal function. *Annu Rev Psychol*. 2010;61(111–40):C1–3.
 125. Tsukahara S, Morishita M. Sexually dimorphic formation of the preoptic area and the bed nucleus of the stria terminalis by neuroestrogens. *Front Neurosci*. 2020;14:797.
 126. Elliott KC, Cheruvelil KS, Montgomery GM, Soranno PA. Conceptions of good science in our data-rich world. *Bioscience*. 2016;66:880–9.
 127. Lai M-C, Szatmari P. Sex and gender impacts on the behavioural presentation and recognition of autism. *Curr Opin Psychiatry*. 2020;33:117–23.
 128. Nebel RA, Aggarwal NT, Barnes LL, Gallagher A, Goldstein JM, Kantarci K, et al. Understanding the impact of sex and gender in Alzheimer's disease: a call to action. *Alzheimers Dement*. 2018;14:1171–83.
 129. Becker JB, McClellan ML, Reed BG. Sex differences, gender and addiction. *J Neurosci Res*. 2017;95:136–47.
 130. Del Giudice M, Lippha RA, Puts DA, Bailey DH, Bailey JM, Schmitt DP. Joel et al.'s method systematically fails to detect large, consistent sex differences. *Proc Natl Acad Sci U S A*. 2016;113:E1965–E1965.
 131. Joel D, Berman Z, Tavor I, Wexler N, Gaber O, Stein Y, et al. Sex beyond the genitalia: the human brain mosaic. *Proc Natl Acad Sci U S A*. 2015;112:15468–73.
 132. Sanchis-Segura C, Aguirre N, Javier Cruz-Gómez A, Félix A, Forn C. Beyond, "Sex Prediction": Estimating and Interpreting Multivariate Sex Differences and Similarities in the Brain. *Res Squ*. 2022; <https://doi.org/10.21203/rs.3.rs-741734/v2>.
 133. Smith SM, Nichols TE. Statistical challenges in "big data" human neuroimaging. *Neuron*. 2018;97:263–8.
 134. Hoops D, Ullmann JFP, Janke AL, Vidal-Garcia M, Stait-Gardner T, Dwihapsari Y, et al. Sexual selection predicts brain structure in dragon lizards. *J Evol Biol*. 2017;30:244–56.
 135. Hoops D, Whiting MJ, Keogh JS. A smaller habenula is associated with increasing intensity of sexual selection. *Brain Behav Evol*. 2022. <https://doi.org/10.1159/000521750>.
 136. Hyde JS. The gender similarities hypothesis. *Am Psychol*. 2005;60:581.
 137. Zell E, Krizan Z, Teeter SR. Evaluating gender similarities and differences using metasynthesis. *Am Psychol*. 2015;70:10–20.
 138. Del Giudice M. On the real magnitude of psychological sex differences. *Evol Psychol*. 2009. <https://doi.org/10.1177/147470490900700209>.
 139. Del Giudice M, Booth T, Irwing P. The distance between Mars and Venus: measuring global sex differences in personality. *PLoS ONE*. 2012;7:e29265.
 140. Bishop DVM, Jacobs PA, Lachlan K, Wellesley D, Barnicoat A, Boyd PA, et al. Autism, language and communication in children with sex chromosome trisomies. *Arch Dis Child*. 2011;96:954–9.
 141. Arnold AP, Chen X. What does the "four core genotypes" mouse model tell us about sex differences in the brain and other tissues? *Front Neuroendocrinol*. 2009;30:1–9.
 142. Jennings CG, Landman R, Zhou Y, Sharma J, Hyman J, Movshon JA, et al. Opportunities and challenges in modeling human brain disorders in transgenic primates. *Nat Neurosci*. 2016;19:1123–30.
 143. Watase K, Zoghbi HY. Modelling brain diseases in mice: the challenges of design and analysis. *Nat Rev Genet*. 2003;4:296–307.
 144. Zeng Y, Lou SL, Liao WB, Jehle R, Kotrschal A. Sexual selection impacts brain anatomy in frogs and toads. *Ecol Evol*. 2016;6:7070–9.
 145. Gur RC, Gur RE. Complementarity of sex differences in brain and behavior: From laterality to multimodal neuroimaging. *J Neurosci Res*. 2017;95:189–99.
 146. Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatom AS, et al. Towards reproducible brain-wide association studies. *Cold Spring Harbor Laboratory*. 2020. <https://doi.org/10.1101/2020.08.21.257758v1>.
 147. Hyde JS. Gender similarities and differences. *Annu Rev Psychol*. 2014;65:373–98.
 148. Cohen A. Surprise! Male and female brains are the same [Internet]. Fast Company. 2021. <https://www.fastcompany.com/90619636/surprise-male-and-female-brains-are-the-same>. Accessed 12 May 2022.
 149. Wierenga LM, Sexton JA, Laake P, Giedd JN, Tamnes CK, Pediatric Imaging NAGS. A key characteristic of sex differences in the developing brain: greater variability in brain structure of boys than girls. *Cereb Cortex*. 2018;28:2741–51.
 150. O'Dea RE, Lagisz M, Jennions MD, Nakagawa S. Gender differences in individual variation in academic grades fail to fit expected patterns for STEM. *Nat Commun*. 2018;9:3777.
 151. Buzsáki G, Logothetis N, Singer W. Scaling brain size, keeping timing: evolutionary preservation of brain rhythms. *Neuron*. 2013;80:751–64.

152. Vijayakumar N, Mills KL, Alexander-Bloch A, Tamnes CK, Whittle S. Structural brain development: a review of methodological approaches and best practices. *Dev Cogn Neurosci*. 2018;33:129–48.
153. Woolston C. Discrimination still plagues science. *Nature*. 2021. <https://doi.org/10.1038/d41586-021-03043-y>.
154. AlShebli B, Makovi K, Rahwan T. Retraction Note: The association between early career informal mentorship in academic collaborations and junior author performance. *Nat Commun*. 2020;11:6446.
155. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*. 2002;15:273–89.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

