

UC Santa Barbara

UC Santa Barbara Previously Published Works

Title

Why and How to Account for Sex and Gender in Brain and Behavioral Research.

Permalink

<https://escholarship.org/uc/item/0bv7s440>

Journal

Journal of Neuroscience, 43(37)

Authors

Eliot, Lise

LeBlanc, Hannah

Maney, Donna

et al.

Publication Date




2023-09-13

DOI

10.1523/JNEUROSCI.0020-23.2023

Peer reviewed

Why and How to Account for Sex and Gender in Brain and Behavioral Research

Lise Eliot,¹  Annaliese K. Beery,²  Emily G. Jacobs,³ Hannah F. LeBlanc,⁴ Donna L. Maney,⁵ and  Margaret M. McCarthy⁶

¹Stanson Toshok Center for Brain Function and Repair, Chicago Medical School, Rosalind Franklin University of Medicine & Science, North Chicago, Illinois 60064, ²Department of Integrative Biology, University of California-Berkeley, Berkeley, California 94720, ³Department of Psychological & Brain Sciences, University of California-Santa Barbara, Santa Barbara, California 93106, ⁴Division of the Humanities & Social Sciences, California Institute of Technology, Pasadena, California 91125, ⁵Department of Psychology, Emory University, Atlanta, Georgia 30322, and ⁶Department of Pharmacology, University of Maryland School of Medicine, Baltimore, Maryland 21201

Long overlooked in neuroscience research, sex and gender are increasingly included as key variables potentially impacting all levels of neurobehavioral analysis. Still, many neuroscientists do not understand the difference between the terms “sex” and “gender,” the complexity and nuance of each, or how to best include them as variables in research designs. This TechSights article outlines rationales for considering the influence of sex and gender across taxa, and provides technical guidance for strengthening the rigor and reproducibility of such analyses. This guidance includes the use of appropriate statistical methods for comparing groups as well as controls for key covariates of sex (e.g., total intracranial volume) and gender (e.g., income, caregiver stress, bias). We also recommend approaches for interpreting and communicating sex- and gender-related findings about the brain, which have often been misconstrued by neuroscientists and the lay public alike.

Key words: sexual dimorphism; gender disparity; testosterone; female inclusion; statistical analysis; anti-sexism

Introduction: Why Account for Sex and Gender?

Sex and gender differences in the brain and behavior are of great interest to society and have the potential to impact the diagnosis and treatment of neuropsychiatric and neurologic disorders. But there remains considerable misunderstanding about the difference in meaning between the two terms, the complexity that arises by treating each as a research variable, and thus, uncertainty about how to detect and interpret sex and gender influences when conducting neuroscientific research.

A consensus report from the National Academies of Sciences, Engineering, and Medicine (2022) has defined “sex” as:

A multidimensional construct based on a cluster of anatomical and physiological traits, that include external genitalia, secondary sex characteristics, gonads, chromosomes, and hormones.

And the report has defined “gender” as:

A multidimensional construct that links **gender identity**, which is a core element of a person’s individual identity;

gender expression, which is how a person signals their gender to others through their behavior and appearance; and cultural expectations about social status, characteristics, and behavior that are associated with sex traits.

Neither variable is strictly binary and each is composed of multiple dimensions that may or may not “match” within a single individual. Further, the two constructs are inextricably linked with each other such that it is almost never possible, particularly in studies of humans, to experimentally separate them. For that reason, in this article we will sometimes refer to them together as “sex/gender” to acknowledge their entanglement.

Many human psychological traits show gender differences, albeit most are minor (Hyde, 2005; Zell et al., 2015). Each of these could be shaped by sex-related factors (e.g., presence or absence of the Y chromosome in mammals), gender-related factors (e.g., occupational experience), or both. This interplay is even more critical when neuroscientists seek to explain gender disparities in mental health and neurologic disorders; diagnoses such as depression, anxiety, eating disorders, and dementia are at least 50% more common in women, whereas ADHD, substance use disorders, dyslexia, and autism spectrum disorders are at least twice as common in men (Fig. 1).

Motivated in part by such disparities, many research funding agencies have begun mandating the consideration of sex and gender in biomedical research. In the United States, this process began with the 1993 NIH Revitalization Act, which required that all clinical trials funded through the agency include women and minority groups when clinically relevant. In neuroscience, gender parity has largely been achieved in imaging studies of the human brain (Eliot et al., 2021) but women continue to be

Received Apr. 18, 2023; revised July 14, 2023; accepted July 18, 2023.

This work was supported by National Institutes of Health Grants R01MH52716 and R01DA039062 to M.M.M., R01AG063843 to E.G.J., and U54AG062334 to D.L.M.; National Science Foundation CAREER 2239635 to A.K.B.; the Fred B. Snite Foundation to L.E.; and the Ann S. Bowers Women’s Brain Health Initiative to E.G.J. This article originated from a professional development workshop at the 2022 Society for Neuroscience annual meeting. We thank Vera Kojcini for her excellent support coordinating this session.

The authors declare no competing financial interests.

Correspondence should be addressed to Lise Eliot at lise.eliot@rosalindfranklin.edu.

<https://doi.org/10.1523/JNEUROSCI.0020-23.2023>

Copyright © 2023 the authors

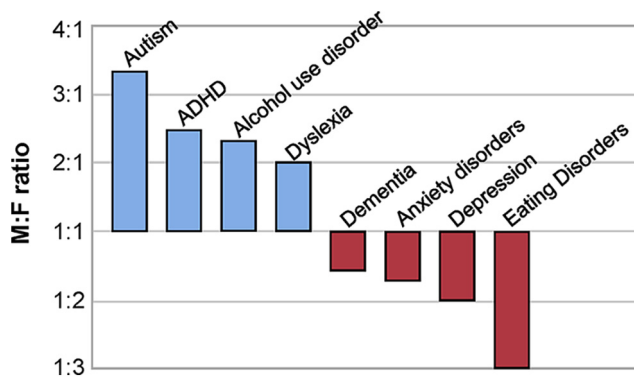


Figure 1. Gender prevalence ratios for common neuropsychiatric disorders. Modified from Eliot et al. (2021). Sources: autism (Loomes et al., 2017); ADHD (Polanczyk et al., 2007); alcohol use disorder (Grant et al., 2004); dyslexia (Rutter et al., 2004); dementia (Buckley et al., 2019); anxiety disorders (McLean et al., 2011); depression (Salk et al., 2017); and eating disorders (Hudson et al., 2007).

under-represented in research on certain disorders, including stroke (Carcel et al., 2021). A greater problem is the failure of researchers to report data disaggregated by sex, even when women and men are included in comparable numbers (Geller et al., 2018). Considering the large sample size of many human studies, this omission amounts to a lost opportunity and short-changes systematic reviews and meta-analyses.

With regard to preclinical research involving animals and cells, it was not until 2016 that the U.S. National Institutes of Health (NIH) mandated that sex be considered, through its “Sex as a Biological Variable” (SABV) policy. In many fields, including neuroscience, female nonhuman animals had been systematically excluded under the mistaken impression that the estrus cycle adds unacceptable variability (Beery and Zucker, 2011). However, meta-analyses of trait variability in male versus unstaged female mice (Prendergast et al., 2014) and rats (Becker et al., 2016) have demonstrated that, while certain traits can be more variable in one sex than another, females of these species are no more variable than males overall and are significantly less variable in spontaneous behavior, both within and across individuals (Levy et al., 2023). Similarly among humans, the coefficient of variation for cognitive, mental, and physical health does not differ between men and women during the reproductively active years (Smarr et al., 2021; Pritschet, 2022). Imaging studies of the brain demonstrate that structural volumes and surface area are modestly more variable in males (Forde et al., 2020; Wierenga et al., 2022). Thus, while pregnancy remains a concern in certain clinical trials (Blehar et al., 2013), in most cases there is no reason to exclude females from basic science or clinical research.

Single-sex studies still have an important place in addressing certain research questions, but for most biomedical topics, including neuroscience, there are ample reasons to include equal numbers of females and males. In addition to exploring the neural correlates of gendered behavioral differences and health disparities, such inclusion can uncover mechanisms of CNS function that may not be evident at the behavioral or even cellular level. Differences in underlying cellular and molecular mechanisms, known as “latent sex differences” (Jain et al., 2019) could potentially explain earlier inconsistencies in the scientific literature when sex was not specified, and theoretically lead to different treatments or diagnostic criteria for boys versus girls or men versus women. Some more prominent examples of latent sex differences in animal studies include the cellular and molecular mediators of pain (Sorge and Strath, 2018), social attachment (De Vries, 2004), reward (Becker and Koob, 2016), and hippocampal endocannabinoid signaling (Tabatadze et al., 2015).

By contrast, other sex effects can be divergent, where a specific context or environmental perturbation invokes a sex difference in behavior from an otherwise similar neurobiological profile. In rodents, for example, stress can have opposite effects on learning and memory in males and females (Shors et al., 2001) and behavioral strategies in response to foot shock can diverge between the sexes (Gruene et al., 2015). In such cases, studies that exclude one sex might provide a distorted view of the phenomenon, and pooling male and female animals may even cancel out an effect of the treatment or exposure (Beery, 2018). The only way to uncover such mechanisms is to include sufficient numbers of male and female animals and to include sex as a factor in statistical analyses.

For United States-based researchers, implementation of the SABV policy and growing awareness of the importance of female inclusion have produced a substantial increase in the proportion of studies including both male and female animals. Over the decade from 2009 to 2019, studies using males and females increased across biological disciplines from 23% to 49% (Beery and Zucker, 2011; Woitowich et al., 2020) (Fig. 2A). But while inclusion has improved, other aspects of reporting and statistical analysis remain problematic. Across 10 domains of preclinical biology using animal subjects, only 50% of mixed-sex studies in 2009 considered sex in their analysis; 10 years later, this number fell to 42% of mixed-sex studies reporting sex-based analyses (Woitowich et al., 2020). In neuroscience, these numbers were 22% and 18%, respectively (Fig. 2B), and only a fraction of those used appropriate analyses to detect sex differences (Garcia-Sifuentes and Maney, 2021). An in-depth analysis of neuroscience and psychiatry papers over the same time interval replicated the finding that most studies did not use an optimal design for discovering sex differences (Rechlin et al., 2022). The authors documented frequent omission of sample size, imbalanced use of males and females, inappropriate statistical approaches, and use of males and females in some but not all parts of a study, revealing considerable room for improvement. Thus, researchers would benefit from greater awareness of the context and components of sex-based analysis to fully achieve the aims of the SABV policy.

For studies of humans, the influence of environmental factors is multiplied many-fold because of the contribution of gender. In human societies, gender roles and other gender-related categorizations can have profound effects on the brain and behavior throughout the lifespan. Thus, in addition to accounting for sex, human neuroscience research must take account of gender; that is, gender norms, power relations, economic security, life experiences, and other factors that may contribute to disparities favoring men or women. Such approaches are now advocated by public funding agencies including the Canadian Institutes of Health and the European Commission (White et al., 2021). Gender analysis is recommended, though not yet required, by the U.S. NIH (U.S. Department of Health and Human Services, 2023).

Given the potential importance of sex and gender to brain and behavioral function, and the opportunities they provide to uncover novel neurobiological processes and foster new avenues for translational research, neuroscientists should invest greater effort in understanding and accounting for their effects. Doing so, however, requires more than simply including equal numbers of male and female subjects. In the remainder of this article, we will address the “how” of studying sex- and gender-related influences on the brain and behavior, including considerations of sample size and statistical comparisons, key covariates of both

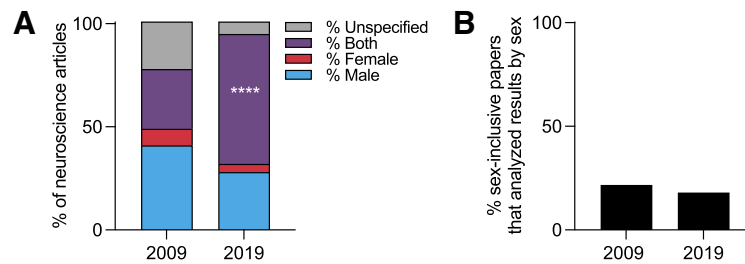


Figure 2. Adoption and analysis of mixed-sex studies show improvement and room for growth. **A**, Across biological subdisciplines and in neuroscience (shown here), studies including males and females in at least one part of the study became significantly more common over a decade. **B**, Despite progress in female inclusion, there was no increase in the percent of studies that conducted sex-based analysis. A substantial proportion of studies that did conduct sex-based analyses used inappropriate statistical approaches (García-Sifuentes and Maney, 2021). **A, B**, Based on data from Weitowich et al. (2020).

sex and gender, strategies for analyzing mechanisms underlying male/female difference, and methods for operationalizing “gender” in human studies. We end with a focus on how sex- and gender-related findings are interpreted and communicated to the nonscientific public, an often-fraught enterprise in need of greater thoughtfulness on the part of researchers (Eliot, 2011; Maney, 2014; Rippon et al., 2021). After each main section, we offer specific recommendations to aid neuroscientists in their reporting and analyses of sex and/or gender effects.

Understanding Sex Effects

Operationalization of sex

Although we often think of sex as a single binary variable, it is actually a complex phenotype composed of many physiological elements that can change dramatically at different stages of development and in different environmental contexts. Thus, because it is not a unitary variable, the first step in accounting for sex is to operationalize it using a stated measure (e.g., karyotype, anogenital distance, self-report) on which data are subsequently collected and published. The measure chosen to represent sex does not have to be the same across laboratories or even across studies within a laboratory; for reasons of practicality, sex is typically operationalized using different measures for cells in culture versus mice ordered from a vendor versus patients in a clinical trial. In any of these cases, the measure used should be transparently reported and consistent within a study (Richardson, 2022).

Once it is decided which variable will represent sex in a study, it is equally important to recognize variables that covary with that one; these additional variables have potential to confound the interpretation of results. In the sections below, we review some of the mechanisms that are typically included in the construct of sex, plus others not typically included that can nonetheless give rise to brain and bodily sex differences. Our goal in these sections is to raise researchers’ awareness of proximal causes of sex differences as well as covariates of sex that could be addressed. Sex differences in the brain and behavior are never a purely direct effect of genes, gonads, or any other variable used to operationalize sex; rather, they involve intermediary steps. In rodents, for example, sex differences can result from differential transcription of genes, apoptosis of cells in specific brain regions, and differential responses of a parent to male and female offspring (McCarthy and Arnold, 2011). So it is typically these underlying, sex-related variables or processes that are informative for elucidating the neural basis of male–female difference, along with phenotypic variance within sex. Sometimes, sex-related variables, such as body size or longevity, may provide simple explanations for apparent neural or clinical differences

between males and females; in other situations, cellular and molecular processes may differ, opening up unique opportunities for advances in basic and clinical science.

Effects of genes

The determination of gonadal phenotype in mammals is controlled by a single gene on the Y chromosome, the Sry gene (sex-determining region of the Y chromosome), which codes for a transcription factor that initiates a gene expression cascade that drives the undifferentiated gonad toward a testis. If Sry is not present, the gonad develops into an ovary (Goodfellow and Lovell-Badge, 1993). The presence of testes and subsequent prenatal testosterone production is the key initiator of masculinization of the brain and behavior. For these reasons, the presence of the Y chromosome is a common sex-related factor used to operationalize sex. But are there genes on the X or Y chromosome that could directly differentiate the brain and/or impact sex differences in behavior? The advent of a genetically modified mouse in which the Sry gene is relocated to an autosome allowed for this question to be asked definitively, by permitting the comparison of WT animals to XX males (i.e., with testis) and XY females (i.e., with ovaries). Known as the 4-core-genotype (Arnold and Burgoyne, 2004), this model is being used to dissect out the relative influence of gonadal steroid hormones and X or Y chromosome genes. For example, certain features of aggressive and parental behaviors in mice, along with vasopressin immunoreactivity in the lateral septum, were found to be influenced by X or Y chromosomes, independent of gonadal sex (Gatewood et al., 2006). Similarly, using this model, susceptibility to an experimental form of autoimmune encephalomyelitis was found to be greater in XX than XY males (Smith-Bouvier et al., 2008), which may be pertinent to the gender difference in multiple sclerosis prevalence. Importantly, effects of X or Y complement are not necessarily directly attributable to genes on one of those chromosomes, as both contain genes with broad epigenetic effects that can alter expression on the autosomes (Arnold and Chen, 2009).

Expression of genes throughout the genome can now be quantified using deep sequencing of the transcriptome from tissue samples or even single cells, revealing unprecedented levels of specification within what were once considered simple phenotypes, such as inhibitory versus excitatory neurons. The importance of transcriptional networks, as opposed to single genes, further expands the complexity of gene regulatory control and modifiability by hormones and other factors, which may or may not be sex-specific (McCarthy, 2020). Such approaches are uncovering highly specific nodes of specialized cells involved in the control of reproductive and other social behaviors (Raam and Hong, 2021; Knoedler et al., 2022). From

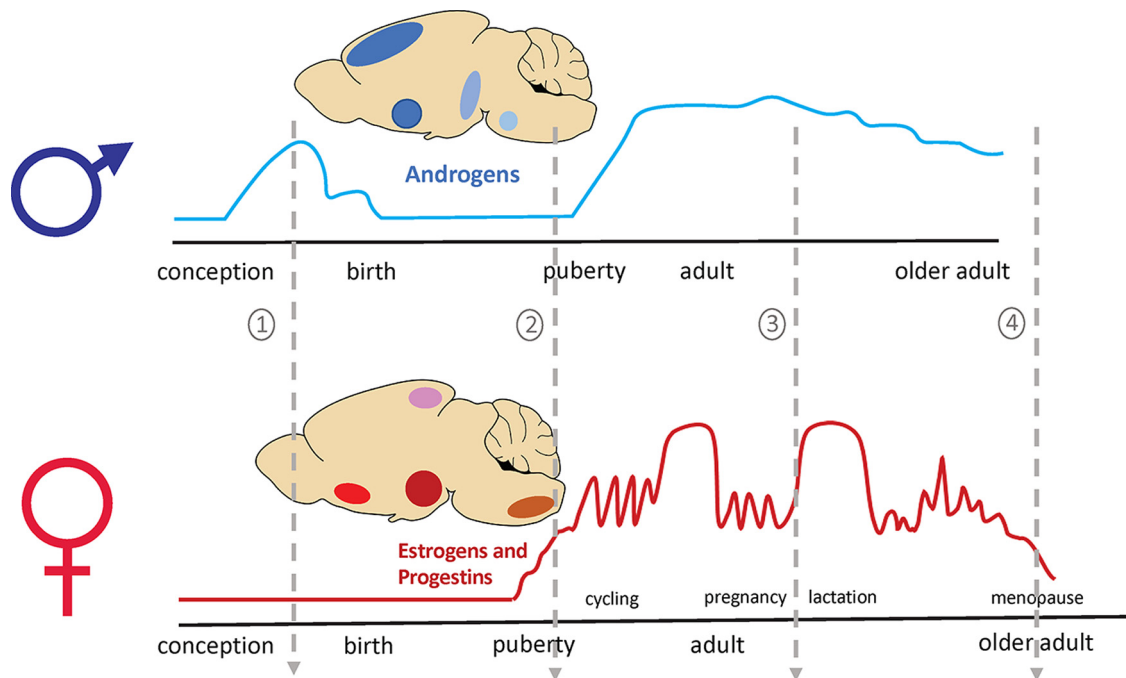


Figure 3. Unique endocrine exposures in male and female rodents over the lifespan. Directly comparing male and female rodents at a given life stage must take into account underlying differences in past or current hormonal milieu and brain maturation. (1) Before birth, males are exposed to relatively high testosterone of their own making, whereas in females, brain exposure to testosterone, estrogens, and progesterins remains low during this time. (2) Females enter puberty sooner than males; and (3) during the reproductive years, females may be cycling, pregnant, or lactating, all of which are physiological states distinct from the relatively stable and high levels of testosterone in males. (4) Later in life, females spend considerable time in a post-reproductive state with very low steroid hormone levels. This figure illustrates general hormonal changes across the rodent lifespan, not precise levels, and omits the exposure of males to estrogens and progesterins and females to androgens. Colored ovals in the brains represent hypothetical regions exhibiting sex-specific characteristics during particular life stages or hormonal conditions. Observations of sex differences are appropriately framed when all of these variables are considered.

a disease perspective, transcriptomics has revealed vast and unpredicted patterns of sex-specific gene expression associated with the same diagnostic phenotype in males and females (Labonté et al., 2017; Massey et al., 2021). However, a word of caution is warranted when unbiased transcriptomic analysis is benchmarked against databases that themselves may be sex-biased. The five most commonly used omics resources are Gene Ontology, KEGG, Reactome, Wiki Pathways, and Panther, from which are derived third-party tools, such as DAVID and Profiler. The omics resources all provide the citations from which the original data were drawn, but none of them annotates the data by sex. Given that the bulk of biomedical research over the last few decades has been conducted on male subjects, there is a strong potential for embedded bias in the databases (Bond et al., 2021). Such bias could manifest as greater accuracy in one sex, missed components in one sex, or frank inaccuracy in one sex. Independent verification, by PCR or other quantitative means, is the only way at present to be assured that latent or hidden biases are not the true origin of a sex difference found using omics databases.

Effects of hormones

Hormones are so closely tied with our conceptualization of sex that they are included in the very definition (e.g., National Academies of Sciences, Engineering, and Medicine, 2022). The hormones most relevant in this context are those produced by the ovary and testis, such as testosterone and estradiol. In the past, these hormones have been referred to as “sex steroids” or “gonadal steroids.” These terms are falling out of use, first because there is no hormone that itself defines a sex; testosterone and estradiol are produced and active in female and male

bodies alike. Further, tissues other than the gonads, such as the brain, can synthesize these steroids *de novo* (Azcoitia et al., 2018). Here, we will call these hormones “steroid hormones” with the caveat that this label can also be used to refer to corticosteroids, sometimes known as stress hormones. We recommend using the specific name of the hormone when possible and relevant (e.g., “testosterone”) or the more general terms “androgens” or “estrogens.”

Steroid hormones are powerful modulators of biological processes at all levels, often independently of other markers of an animal’s sex (McCarthy and Konkle, 2005). Nonetheless, because plasma levels of androgens and estrogens can differ dramatically between males and females, they are often the default hypothesized mechanism for findings of sex differences (Maney et al., 2023), particularly when the hypothesizing is done *post hoc*. However, steroid hormones are just one of many potential contributors to sex differences. Care should be taken to consider less obvious factors, some of which are discussed below in “Effects of age,” “Effects of body size,” and “Effects of environment.”

Prenatal and neonatal hormones

In rodents, many sex differences in the brain and subsequent behavior arise from differences in prenatal or perinatal exposure to testosterone and its major metabolite, estradiol (Fig. 3). In both rats and mice, the fetal testis begins copious androgen production during late gestation, 3–5 d before birth, and continuing through birth but dropping precipitously thereafter and staying low until puberty. In females, the ovaries remain quiescent, and there is far less exposure to steroids developmentally. Much of what we understand about steroid-mediated sexual differentiation of the brain comes from the fortuitous fact that newborn female rodents

remain sensitive to the actions of testosterone; if it is administered exogenously, the process of masculinization is initiated and the endpoints that are normally achieved in males prenatally can now be triggered in females postnatally, making them an interesting model in which to test the effects of androgens across the lifespan (McCarthy et al., 2018). In humans and nonhuman primates, steroid-mediated brain sexual differentiation begins during the second trimester, and this surge is largely complete by birth, greatly challenging the ability to interrogate mechanisms. Regardless, some behaviors, along with the sensitivity to steroid hormones in adulthood, appear to be influenced by this prenatal exposure in primates (Breedlove, 1994; Wallen and Baum, 2002), making it an important period in brain sexual differentiation.

Rhythmic changes in hormone production

A central feature of the mammalian endocrine system is that hormone production varies over time. Many of these endogenous endocrine rhythms occur independently of sex. For example, gonadotropin and steroid hormones exhibit circadian rhythms tied to the sleep-wake cycle in males and females. Testosterone and cortisol production peak in the A.M. and decline in the P.M., a circadian pattern that is common to males and females. In other instances, the frequency, duration, and intensity of these endocrine rhythms can differ by sex. For example, ovarian cycles, such as the 4- to 5-d estrous cycle in rodents and ~28 d menstrual cycle in humans, produce cyclic changes in circulating estrogens and progesterone. During pregnancy, serum estrogens and progesterone concentrations increase ~100-fold and then decline precipitously at parturition. Each of these neuroendocrine transitions drives physiological changes throughout the body, in some cases including the brain. For example, in a series of foundational studies in rats, Woolley et al. (1990) identified changes in hippocampal spine density in CA1 neurons that are tied to stage of the estrous cycle, with a ~30% increase in spines during proestrus (the day of ovulation with estradiol peaks) relative to estrus 24 h later. Similarly rapid effects of steroid hormones on hippocampal morphology and the functional connectome have been observed over the human menstrual cycle (Pritschet et al., 2020, 2021; Taylor et al., 2020), with ultra high-field imaging recently revealing volumetric increases in hippocampal CA1 during high estradiol/low progesterone phases (Zsido et al., 2023). During pregnancy, organizational effects of steroid hormones are likely responsible for brain morphologic changes that have been observed between prepartum and postpartum phases in women (Hoekzema et al., 2017). Neuroanatomical adaptations are also evident in new fathers during the transition to parenthood (Martínez-García et al., 2023), although their endocrine drivers and time course likely differ from the neuronal changes observed in mothers.

Effects of age

Depending on species, males and females may develop, mature, and age at different rates. In humans, the sex difference in longevity is one of the more robust findings, not seen in many other mammals (Austad and Fischer, 2016). Although its basis is not well understood, this longevity gap in humans has a major impact on neurodegenerative disorders, notably Alzheimer's disease, where it largely explains the nearly double number of women as men (Fig. 1) living with the disease (Fiest et al., 2016; Buckley et al., 2019). Maturational sex differences are also important earlier in life, as girls undergo puberty and reach their peak brain volume, gray matter density, and cortical surface area 1-

2 years earlier than boys (Kaczurkin et al., 2019). Such asynchrony is pertinent when comparing adolescent brain or behavioral measures between sexes, since what looks like a sex difference at a particular age may be better explained by a difference in pubertal timing. Conversely, when comparing boys and girls at similar pubertal stages, their mismatch in age may explain any brain or behavioral differences better than differences in steroid hormone levels (Vijayakumar et al., 2018).

Indeed, age is often an important variable when considering the effects of hormones. In rodents, marked differences in hormonal profiles of males and females occur across the lifespan, beginning before birth, diverging at puberty, and then again after reproductive senescence or menopause in females. Moreover, females experience unique hormonal fluctuations (e.g., the estrus cycle, pregnancy), in addition to the diurnal changes in hormone production experienced by both males and females. This variation presents a challenge when one is directly comparing males and females as they may differ by both maturational status and hormonal profile (Fig. 3). There is no simple answer to this conundrum, but awareness of what is actually being compared is essential. For example, an insult to the nervous system in a 35-d-old mouse or rat will coincide with a prepubertal window in males but a postpubertal window in females, because of differences in the timing of the onset of puberty. Likewise, aged males are likely to have circulating testosterone similar to that of a younger adult, whereas aged females will have minimal circulating estrogens or progestins.

In humans, reproductive aging (defined as changes in steroid hormone production that occur after mid-life) is a major contributor to physiological changes in both men and women. For men, age-related changes in steroid hormone production are linear and protracted, gradually declining beginning in the mid-30s. In contrast, women undergo a more complex transition at mid-life, marked by high fluctuations in steroid hormone production during the perimenopausal phase, and culminating in ovarian senescence at the end of menopause, with the median age of complete reproductive senescence at 52.4 years in the United States (Gold et al., 2013). The menopausal transition results in a substantial decline in ovarian hormone production: up to 90% for both estradiol and progesterone. Human studies of the aging brain should therefore distinguish between the effects of chronological versus endocrine aging, although many do not. One approach is to capitalize on variation in the timing of the menopausal transition, matching participants on chronological age while allowing reproductive stage to vary (Jacobs et al., 2016). Another approach is to study the effects of pharmacological manipulations that temporarily and reversibly induce a menopause-like state (Berman et al., 1997).

Effects of body size

In many of the commonly studied mammals, males have larger bodies than females. While there are many sex differences that size cannot account for, body weight makes an important contribution to many traits (Wilson et al., 2022). This size difference is reflected in brain volumes, with differences ranging from 2.5% larger male brain volume in C57BL/6J mice (Spring et al., 2007) to 10% larger in Fischer 344 rats (Goerzen et al., 2020) and 11% larger in humans (Williams et al., 2021). For studies exploring regional brain differences between males and females, it is therefore necessary to control for total brain volume (TBV) or intracranial volume (ICV), which largely eliminates regional volume differences in humans (Jäncke et al., 2015; Ritchie et al., 2018). The small (1%-2%) volumetric

sex differences that remain after controlling for TBV or ICV have proven quite variable and dependent on image processing pipeline and the method of size normalization (Eliot et al., 2021). Absolute brain size also affects other features of brain architecture, such as the ratio of white:gray matter and the ratio of interhemispheric to intrahemispheric connections. Larger brains need more or thicker white matter pathways to interconnect more distant regions, and interhemispheric traffic grows inefficient with increases in brain size (Zhang and Sejnowski, 2000). Hence, men's brains have ~6% higher white:gray matter ratio (Pintzka et al., 2015; Ritchie et al., 2018) and a lower ratio of interhemispheric to intrahemispheric connections (Ingalhalikar et al., 2014) than women's brains, both of which are attributable to brain size, not sex per se (Lewis et al., 2009; Hänggi et al., 2014). In other words, such measures differ between large- and small-headed men as much as between men and women, so are unlikely relevant to behavioral or clinical gender differences. These findings highlight the importance of controlling for individual ICV or TBV when conducting any study of individual brain structural difference, gender-related or otherwise (Jäncke et al., 2015), using correction factors that may differ for different brain compartments (Jäncke, 2018; Williams et al., 2021).

Effects of environment

Still other important, but minimally studied, contributors to brain sex differences are social and environmental factors. Although the term “gender” is usually restricted to humans, male and female animals often inhabit different social spaces and are faced with different experiences that can alter their brains and behavior. Such differential experience can begin even before birth; for example, females' exposure to a male co-twin or adjacent male littermate may shift certain features of physiology and behavior in a male-typical direction, presumably as a result of prenatal testosterone exposure (Ryan and Vandenberg, 2002; Tapp et al., 2011).

After birth, the opportunity for sex- or gender-differentiated experience is magnified many-fold. Although we tend to focus on such differential experiences primarily in humans, the social environment can be highly differentiating in other animals as well. For example, rat dams treat male and female pups differently from the first days of life, licking and grooming males more than females, a difference that acts to shape the neural circuit underlying male copulatory behavior (Moore, 1992). Since maternal licking-and-grooming, acting through epigenetic mechanisms, is also known to influence pups' stress regulation, exploratory behavior, and spatial learning (Champagne and Curley, 2009), such sex-specific rearing differences could contribute to brain and behavioral sex differences. In nonhuman primates, too, the complex social environment is known to foster sex differences in behavior; for example, wild chimpanzees dig for termites with flexible sticks, a useful skill that females in some locales learn earlier than males, because young females spend more time watching their mothers “fish” in this way (Lonsdorf et al., 2004).

One environmental factor that has been widely studied is stress, which can produce effects that sometimes, but not always, depend on sex (McCarthy, 2016). For example, Bohacek et al. (2015) found sex differences in stress-induced hippocampal gene expression in response to swim stress, but not restraint stress. Research in humans and nonhuman animals has now documented the impact of a variety of stressors, occurring at various times across the lifespan, on the emergence of brain and behavioral sex differences, especially in psychopathology (Bale and

Epperson, 2015). Such findings illustrate the importance of environmental context in generating neurobehavioral sex differences, and may be especially impactful given the different stressors that are often exerted on male versus female bodies (Hyde and Mezulis, 2020).

Recommendations for studying sex effects

An updated understanding of sex as a multidimensional construct leads to several recommendations for analyzing its influence in brain and behavioral studies:

- Clearly operationalize sex using a stated measure and collect data on that measure.
- Recognize that factors contributing to sex differences include genes, hormones, developmental stage, and body size.
- When conducting transcriptomic analyses using so-called “unbiased” approaches, such as RNA-Seq, be aware of potential hidden bias in the publicly available omics databases and independently verify differences you consider important.
- Avoid the term “sex hormones” since the actions of these agents are not limited by sex.
- Acknowledge and assess, when possible, the broad environmental factors that can lead to differences between and among males and females, including intrauterine exposures, parental care, social grouping, stressors, and other experiences across the lifespan.

Understanding Gender Effects

This brings us to a preeminent part of human experience that has thus far been largely ignored in our effort to understand male-female brain differences: the impact of gender. As defined above, gender is a complex psychosocial construct that can have profound effects on experiences and, therefore, the brain. In some cases, gender acts independently of sex, and in others, sex and gender interact over the lifespan (Krieger, 2003; Springer et al., 2012; Hyde et al., 2019).

Researchers outside of neuroscience have found that sex and gender can have independent effects on health and can do so even in cases where there is no obvious difference between women and men. For example, among younger patients diagnosed with acute coronary syndrome, Pelletier et al. (2016) observed no difference in the risk of recurrence between women and men, but instead found that patients with a higher “femininity score” were more likely to experience a recurrence of acute coronary syndrome, regardless of their sex.

Operationalizing gender as separate from sex shows promise in brain imaging studies as well. In a recent review, Rauch and Eliot (2022) identified eight neuroimaging studies that have assessed brain parameters using gender as an independent variable. While still preliminary, this work provides evidence that gender attributes can explain some of the individual variance in regional brain structure and functional circuitry, apart from or in addition to binary sex effects.

Despite these encouraging efforts, research using gender as a variable of interest is still in its infancy, in part because of the inadequacy of existing instruments for its operationalization. Existing measures largely fail to capture gender's nonbinary, multidimensional nature (Horstmann et al., 2022). Research that operationalizes gender as either one-dimensional (a bipolar continuum that ranges from “masculine” to “feminine”) or two-dimensional (with masculinity and femininity as distinct

spectra) cannot identify the specific gendered factors that affect health or the brain (Nielsen et al., 2021; Horstmann et al., 2022). Moreover, such polar conceptualizations of gender often exclude the experiences of trans, nonbinary, and genderqueer individuals (Hyde et al., 2019; Miani et al., 2021).

In addition to being nonbinary, gender operates on many scales: from individuals' gender identities, to the structure of labor markets, to expectations around caregiving work, to experiences of discrimination (Bolte et al., 2021; Miani et al., 2021). While these different dimensions often correlate with one another, any single dimension (e.g., gender identity) may be a poor proxy for any other dimension (e.g., the number of hours spent caring for a family member). Gender remains a primary category of social organization, but any individual may exhibit a unique constellation of such factors, which also intersect with other axes of social inequality in ways that can affect brain health.

To address these problems and create a better tool for operationalizing gender, an international team recently developed a new instrument, called the Stanford Gender-related Variables for Health Research (GVHR) (Nielsen et al., 2021). The GVHR conceptualizes gender across three domains: gender norms, gender-related traits, and gender relations. Gender norms are spoken and unspoken rules enforced through social institutions, such as the workplace, family, and broader culture. Gender-related traits refer to individuals' behaviors and tendencies, which may or may not conform to their gender norms. Gender relations refer to the social relationships between individuals, whose own gender identities and related traits operate within a system of gender norms. Gender norms, traits, and relations are culture- and context-specific and change over time.

The GVHR is a questionnaire of 25 items that measure seven gender-related variables across these three domains. The variables are as follows: caregiver strain and work strain (gender norms); independence, risk-taking, and emotional intelligence (gender-related traits); and social support and discrimination (gender relations). The questionnaire was validated in three online samples totaling >4000 diverse respondents.

The GVHR has several advantages for neuroscience researchers. First, the scales make no assumptions about whether traits are "masculine" or "feminine," as such labels are often normative and imprecise. Second, each of the variables is scored separately, rather than being collapsed into a single gender metric. This scoring system improves the precision of the measure, gives researchers the opportunity to select variables that are most relevant to their scientific question, and makes it appropriate for use with nonbinary and trans populations. The GVHR does not replace the independent variable of gender identity (e.g., woman, man, genderqueer, genderfluid, etc.), but may be used in combination with it to better characterize participants within the gender space. Using the GVHR or similar instruments in combination with large-scale neuroimaging should enhance our understanding of the neural basis of gender-related behaviors and psychopathologies.

Recommendations for studying gender effects

Growing awareness of gender as a multidimensional variable with biomedical impact should broaden the way neuroscientists study the human brain and behavior, including the following:

- Whenever possible in human studies, operationalize gender as a variable separate from sex.
- Appreciate the key components of gender beyond identity that may impact brain health and incorporate such measures as appropriate.

Analytic Approaches

Whether one is studying animals or humans, accounting for sex or gender adds complexity to statistical analyses. In this section, we address issues of sample size and analytical approaches. In alignment with the existing SABV policy, we focus on analyses of binary group differences, recognizing that more complex statistical models are needed for analyses that use continuous measures of sex or gender, such as the GVHR, as the independent variable.

Design considerations

Mere inclusion of females and males is not enough to provide insight into the impact of sex/gender on neurobehavioral phenomena. Having rejected concerns that females are more variable than males (Prendergast et al., 2014; Becker et al., 2016), we must still consider the possibility of differences between males and females and their potential influence on other findings.

When sex-inclusive data are analyzed without sex as a factor, sex differences have the potential to increase the variability of the sample as a whole. Fortunately, inclusion of sex as a factor in analysis recovers statistical power in many instances, whether sex differences are present or absent (Fig. 4) (see also Phillips et al., 2023). Thus, to harvest the "low hanging fruit," researchers should analyze data from males and females with sex (formally operationalized) as a factor in the analysis and appropriately balance sample sizes. Of note, some studies treat sex/gender as a nuisance variable in their models, accounting or "controlling" for it without noting possible effects (Beltz et al., 2019). Whether or not sex contributes to the final results, the presence or absence of sex differences and interactions between sex and treatment should be reported, descriptive data should be disaggregated and reported by sex, and consideration should be given to the statistical power of the dataset to detect sex differences.

Exploratory versus confirmatory studies

In their guidance on how to incorporate SABV into research designs and analysis, NIH distinguishes between studies "intended" and "not intended" to detect sex differences (National Institutes of Health, 2015). In the former scenario, testing for sex differences is an *a priori*, stated goal of the study and included among its hypotheses. This sort of research, which we will call "confirmatory" (as opposed to "exploratory," see below), must be powered to test hypotheses about sex. Depending on the question, the sexes might simply be compared with each and no other factors, such as a drug treatment, need be considered. In most cases, however, authors may be testing the extent to which males and females respond differently to a treatment or exposure. For such "factorial" designs, one can test for effects of one factor (e.g., treatment) and effects of another factor (e.g., sex) in the same statistical model. These "main effects" can indicate whether the treatment had an effect overall, regardless of sex (Fig. 4A) and whether there was a sex difference in the outcome measure regardless of treatment (Fig. 4B). However, a main effect of sex does not indicate that females and males responded differently to the treatment. For example, in a study examining the effect of an intervention on stress hormone levels, a main effect of sex would indicate a significant sex difference in stress hormone levels but does not provide evidence that males and females responded differently to the intervention.

For studies with factorial designs, testing whether males and females responded differently to a drug, intervention, or exposure can be done by including an interaction term in the statistical

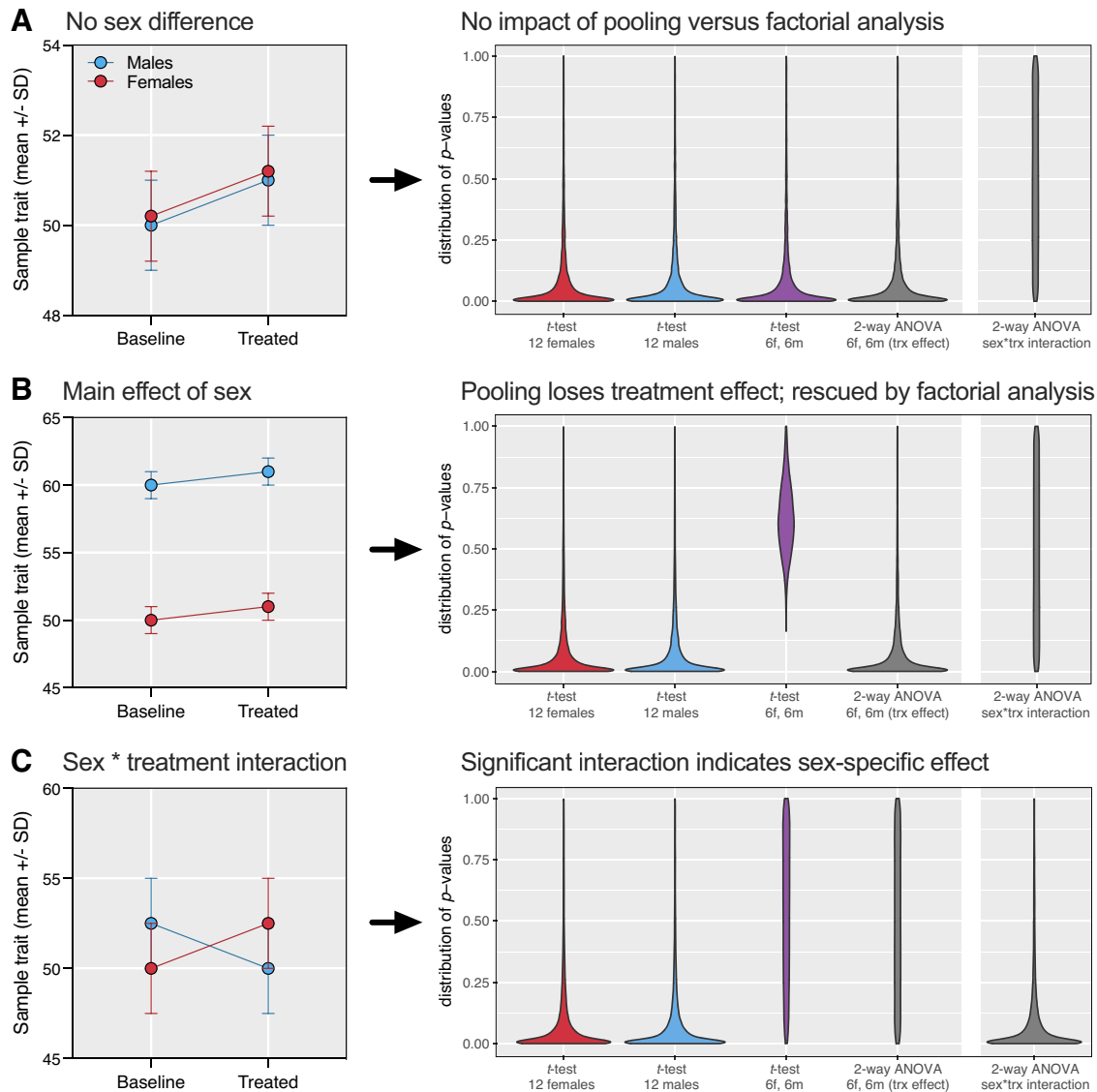


Figure 4. Example data scenarios and impact of analysis as single-sex samples (only females or only males tested), pooled mixed-sex samples, or factorial analysis of mixed-sex samples (two-way ANOVA assessing effects of sex, treatment, and their interaction). Left panels, Mean and SD of each group (12 females or males in single-sex studies, or 6 males and 6 females in mixed-sex groups). Effect sizes in each sex were matched (Cohen’s $d = 1.0$) and 10,000 simulated groups of each type were generated for comparisons. Right panels, Statistical consequences of different analysis paths as violin plots of p values generated for each test across 10,000 samples. Distributions weighted closer to zero (e.g., red distributions) yielded “significant” treatment differences in more simulated comparisons. Red represents effect of treatment across all-female comparisons. Blue represents effect of treatment in all-male comparisons. Purple represents effect of treatment when males and females are pooled and sample size is not increased. Gray represents treatment and sex \times interaction effects in a two-way ANOVA. **A**, Pooling has no impact in the absence of sex differences, but **(B)** obscures the treatment effect in the presence of a main effect of sex, or **(C)** when there is an interaction between sex and treatment (i.e., the treatment affected the sexes differently). Factorial analysis rescues the treatment signal when there is a main effect of sex **(B)**, and reveals a potential sex \times treatment interaction **(C)**. If the sample size is not increased for mixed-sex exploratory studies, a finding such as **C** would call for a future follow-up with a larger sample of males and females, and potential future studies of mechanisms in each sex. Redrawn from and based on simulations performed by Beery (2018).

model, for example, sex \times treatment (Fig. 4C). A significant interaction would indicate that males and females did indeed respond differently to the treatment. Such interactions can require substantial power to detect. Thus, studies intended to reveal “sex-specific” effects of another factor typically require a larger sample size than is required to detect the factor’s effect when the sexes are pooled into a single sample. Other authors have provided more detailed guidance on determining sample sizes when the goal is to detect a sex difference in the effect of another factor (e.g., Buch et al., 2019; Galea et al., 2020; Phillips et al., 2023).

Sometimes, researchers are not focused on sex differences: they are including males and females to increase the generalizability of their findings, or to make their work more inclusive, or simply to comply with the SABV policy. In these “exploratory”

studies, per NIH guidance, researchers need not increase the sample size beyond that necessary to detect the main effects of other variables of interest, such as exposure, time, or treatment (National Institutes of Health, 2020). Most main effects that can be detected with a certain number of men or males, for example, should theoretically be detectable using a similarly sized sample of mixed sex or gender (Buch et al., 2019; National Institutes of Health, 2020; Phillips et al., 2023). Notably, NIH expects researchers to consider sex even in exploratory studies that are not powered to compare the sexes (National Institutes of Health 2015, 2020; Clayton and Tannenbaum, 2016; Cornelison and Clayton, 2017; Clayton, 2018). However, the latter practice can potentially lead to inaccurate pronouncements of sex or gender differences (false

positives); and conversely, real sex or gender differences may not be detectable with small sample sizes, leading to the reporting of false negatives. In the next sections, we discuss some common mistakes and how to avoid them.

How to test for a sex-specific effect

NIH recommends that researchers analyze and report data separately for males and females (Clayton and Tannenbaum, 2016; Cornelison and Clayton, 2017; Clayton, 2018; National Institutes of Health, 2020). Many journals offer similar guidance (Heidari et al., 2016). Although this reporting is important to promote transparency and future meta-analyses, the analysis of disaggregated data can lead to a well-known analytical error. If two independent samples are analyzed separately and a significant effect is found in one but not the other (e.g., $p = 0.02$ vs. 0.20), such a result does not constitute evidence that the effect in one sample is reliably larger or smaller than in the other (Gelman and Stern, 2006; Sainani, 2010; Nieuwenhuis et al., 2011; Makin and de Xivry, 2019). The problem is that this approach does not test whether an effect in one group is stronger than in another. To show that an effect is significantly different between two sexes, we must directly compare the two effects by testing for an interaction between the factor in question and sex (Gelman and Stern, 2006; Sainani, 2010; Nieuwenhuis et al., 2011; Bland and Altman, 2015; George et al., 2016; Makin and de Xivry, 2019; Vorland, 2021; Vorland et al., 2021).

Failing to test formally for sex-specific effects is so rampant an error that it has earned its own acronym: the Difference in Sex-Specific Significance (DISS) error (Maney and Rich-Edwards, 2023). DISS errors most often result in Type I errors (Bland and Altman, 2015), biasing researchers toward positive findings. But DISS approaches can also lead to Type II errors, or false negatives, where sex-specific effects are missed (Vorland et al., 2023). Both types of errors hamper reproducibility.

When assessing sex-specific effects, it is therefore critical to test for interactions between sex and other factors of interest (e.g., Fig. 4, gray plots) and not simply compare the statistical significance of that effect within each sex. Testing for interactions will increase power to detect effects of the other factors (because all of the samples can be considered), while at the same time ensuring that exploratory testing for potential sex differences is done in a statistically valid way. This rigorous approach is especially important in studies not powered to detect small- or medium-sized interactions, since the false positive rate of the DISS error can reach 50% in such cases (Allison et al., 2016). If the interaction with sex is significant, then one can follow-up with *post hoc* tests within each sex. If the interaction is not significant, however, the *post hoc* tests are not warranted and should not be performed; authors should instead report that there is no evidence that the sexes responded differently, with the caveat that the study may be underpowered. In these cases, it may be appropriate to call for follow-up studies to test for sex-specific effects of the treatment or exposure.

Before statistically comparing the sexes, it is important to discern whether the sexes are actually in a comparable state. In the “Effects of age” section above, we pointed out that same-age males and females can be at very different stages of pubertal development or reproductive senescence. In such cases, when sex is strongly confounded with another relevant variable, it may be appropriate to test hypotheses only within-sex rather than including males and females in the same statistical model. This approach does not compare the sexes with each other; thus, no claim can be made about sex difference per se.

Regardless of the analytical approach, whenever mixed-sex studies are performed, it is imperative to make all raw data, disaggregated by sex, available for future investigations. This can be done either in a main paper or in its Supplemental Material, and will facilitate subsequent meta-analyses and help avoid “file drawer” omissions (Rosenthal, 1979).

Recommendations

To comply with sex/gender inclusion guidelines and avoid common statistical errors, we offer the following recommendations:

- Balance female and male sample sizes as appropriate for the research question (except in cases when single-sex studies can be justified).
- Analyze results with sex/gender as an explicit factor in the statistical model.
- Report all statistics relevant to sex comparisons, including main effects and interactions.
- For exploratory, underpowered studies, acknowledge the limitations of the approach without relaxing statistical rigor.
- Regardless of the outcome of sex comparison, publish the sex-disaggregated data for each sample either in the main paper or supplemental material.

Interpretation and Communication

Because the framing of male/female brain differences has broad impact on issues related to gender equity, health disparities, and the rights of sexual and gender minorities, we end with guidance on the interpretation and communication of such differences. Neuroscientific findings hold great appeal in the public imagination, particularly as they relate to differences between men and women or boys and girls. But too often, complex and unreplicated findings about male/female brain differences are simplified and overhyped by popular media, propagating stereotypes and distorting the real science they are purporting to convey (Eliot, 2011; Maney, 2014, 2016; Rippon et al., 2021). When it comes to research on sex and gender difference, neuroscientists should therefore pay closer attention to: (1) negative findings, which often do not get mentioned in studies including both males and females; (2) the actual magnitude (both raw measure and effect size) of group-level sex/gender differences, as opposed to mere statistical significance; (3) the practical relevance of such findings to biomedical treatment or policy-making; and (4) the multifactorial determinants (e.g., genetic, morphologic, psychological, social) that may underlie such differences, before attempting to speculate on their clinical or behavioral importance. Just as the National Academies of Sciences, Engineering, and Medicine (2023) recently warned against “typological thinking” about race and ancestry in the dissemination of genetics and genomics research, sex/gender difference researchers should be cautious about interpreting group-level differences as fixed and attributable to “biological” processes, especially when the underlying mechanisms are largely unknown.

As a baseline, neuroscientists are advised to avoid the term “dimorphism” to describe statistical differences between male and female brains (McCarthy et al., 2012), which rarely satisfy the proper definition of “two forms.” The term “dimorphism” refers to two distinct types, or distinct structures present in only one, but not the other sex. To date, dimorphic structures have not been found in the central nervous systems of vertebrates, with the possible exceptions of striatal Area X in some songbirds (Nottebohm and Arnold, 1976) and the bulbocavernosus spinal motor nucleus in some mammals (Breedlove and

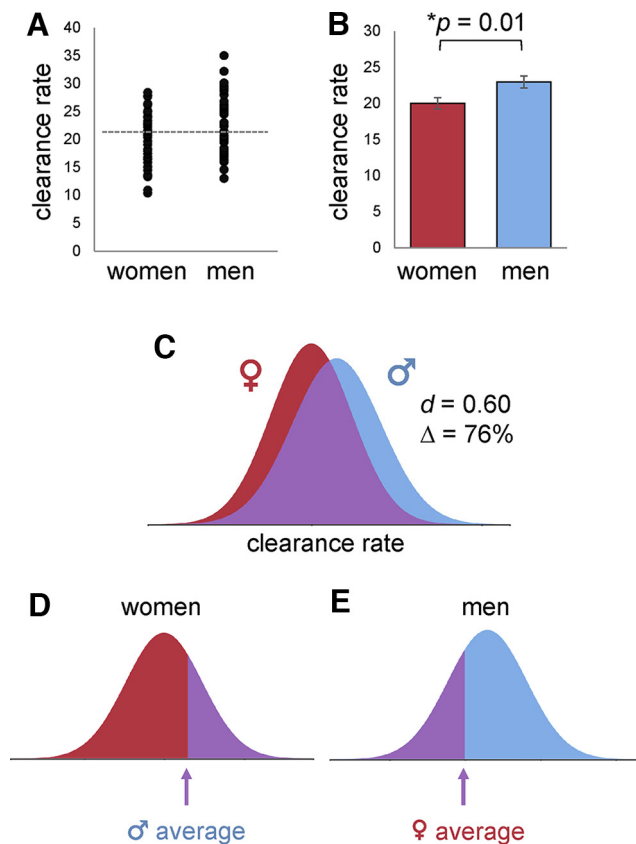


Figure 5. Importance of considering variance within sex/gender and overlap between them. **A**, Hypothetical raw data for the clearance of a drug. **B**, Means and SEM in men and women. A t -test produces a p value of 0.01, meaning that this is a significant sex difference. The effect size, depicted as Cohen's d (**C**), is 0.60, a medium-sized effect that warrants attention (Klein et al., 2015). But despite the low p value and compelling effect size, closer inspection reveals that there are about the same number of men and women above and below the mean (22 men and 18 women above the dotted line, and 18 men and 22 women below). The overlap between the sexes, depicted as Weitzman's δ , is 76%. If we were to guess the sex of one of these participants knowing only their clearance rate, our accuracy would be little better than 50%. Further, (**D**) one-third of the men have a clearance rate lower than the average woman and 30% of the women are clearing the drug faster than the average man (**E**). Sex-specific dosing of this drug, therefore, might benefit patients in the tails of the distribution, but most patients could get a suboptimal dose. For the dataset and guidance on how to create graphs like **C**, see Maney (2016).

Arnold, 1981). Like racial categorization, the term “dimorphism” promotes typological or essentialist thinking, whereas most brain features are more appropriately described as “monomorphic” across the sexes of humans and common experimental animals (Eliot et al., 2021). Indeed, for most measures of mammalian brains and behavior, there is considerable overlap between groups, such that the sex of any individual cannot be identified on the basis of that measure alone. Similarly, a sex difference requiring hundreds of individuals to detect cannot be used to predict the best course of treatment for any single individual (Fig. 5).

As the foregoing discussion makes clear, sex and gender are pertinent to many issues concerning the brain and behavior. However, sex and gender themselves comprise many components and covariates that account for the differences between groups of males and females or variance across all genders (Richardson, 2022). Thus, precise operationalization and attention to confounds is key, and as researchers increasingly incorporate comparisons of males and females, it is important to note this nuance and evaluate these confounds accordingly. For

example, if a certain drug induces more side effects in women than men, should the dosing be adjusted by sex or by body weight, fat-free mass, circulating steroid hormone levels, or some other continuous physiological variable that may better explain variation (Özdemir et al., 2022)? Similarly, when differences are found in neuropsychiatric disease risk, might they be explained by covariates of gender or sex, such as caregiver stress or occupational attainment (Mielke et al., 2018)? These questions are challenging to answer, but vital to address for any drug, treatment, or condition that shows group-level differences between men and women, or boys and girls.

Finally, it is important for anyone commenting on brain sex/gender differences to be mindful of the sexist history that has prefaced research on this topic. Across the many centuries that philosophers and scientists have been addressing brain sex/gender differences, their search has largely focused on the neural basis of women's presumed intellectual and creative inferiority (Schiebinger, 1993; Saini, 2017). Although recent peer-reviewed literature may contain less overt misogyny, the question of brain sex/gender differences is often still posed within a stereotyped frame, for example, referring to men's propensity for “action” versus women's “intuition” (Ingahlalikar et al., 2014). These assumptions demonstrably influence public understanding of gender and reinforce benevolent sexism (O'Connor and Joffe, 2014). Most of this concern is focused on human studies, but as preclinical studies are funded on the basis of their relevance to human health, such considerations are also important when translating findings from nonhuman animals.

As part of an antisexist framework, there is also a renewed call to look beyond sex differences, to study key aspects of the human condition largely ignored by prior generations of research. Neuroscientists know little of how menopause, pregnancy, the menstrual cycle, and hormone-based medications influence the brain (Taylor et al., 2021). Even less of the research on these topics has addressed people across diverse ethnic, socioeconomic, and geographic origins (Petersen et al., 2023). These intersectional blind spots have limited progress on human health generally, and brain health in particular (Taylor et al., 2019).

Thus, special caution applies to the interpretation of neuroscientific findings based on the predominantly White and (presumed) cisgender populations that comprise most of the large and highly used brain/behavioral databases. Available bodies of knowledge largely neglect trans and gender-diverse populations, for example. As we seek to understand the causes of sex/gender behavioral and mental health disparities, it is important to ask, “which men?” and “which women?” and to appreciate that sex and gender lie within an intersectional framework of social identities, all of which interact to shape neural function and behavior (Duchesne and Kaiser Trujillo, 2021). Merely averaging brain measures in one population, no matter how large, may not lead to a better understanding of average male-female behavioral differences and risks endorsing essentialist interpretations that are already the default among many clinicians and the lay public. Little attention has been paid thus far to cross-cultural or ethnic variation in gender expression, to understanding the development and environmental modulation of gender expression, or—more broadly—to understanding the neural basis of gender in any nonbinary way (Rauch and Eliot, 2022). The danger is especially great at present, as legislative bodies increasingly co-opt biomedical findings to define individual rights based on “genetics,” a “person's chromosomes,” or “anatomy” as assessed at birth. By embracing, rather than oversimplifying, the nuance in sex- and gender-related neurobehavioral data, neuroscientists

have the opportunity to challenge such essentialist interpretations that are presently aimed at restricting the rights of sexual and gender minorities (Sudai et al., 2022).

Recommendations

As general guidelines for the reporting and interpretation of sex- and gender-related neurobehavioral data, we recommend that neuroscientists:

- Report negative as well as positive findings.
- Consider effect sizes and practical relevance of sex/gender differences.
- Consider environmental contributions to sex/gender differences, even when the outcome measures are physiological in nature.
- Remain mindful of researcher bias and the sexist history of sex/gender difference research.
- Address the degree to which the study population is representative of human diversity.

References

- Allison DB, Brown AW, George BJ, Kaiser KA (2016) Reproducibility: a tragedy of errors. *Nature* 530:27–29.
- Austad SN, Fischer KE (2016) Sex differences in lifespan. *Cell Metab* 23:1022–1033.
- Arnold AP, Burgoyne PS (2004) Are XX and XY brain cells intrinsically different? *Trends Endocrinol Metab* 15:6–11.
- Arnold AP, Chen X (2009) What does the ‘four core genotypes’ mouse model tell us about sex differences in the brain and other tissues? *Front Neuroendocrinol* 30:1–9.
- Azcoitia I, Arevalo MA, Garcia-Segura LM (2018) Neural-derived estradiol regulates brain plasticity. *J Chem Neuroanat* 89:53–59.
- Bale TL, Epperson CN (2015) Sex differences and stress across the lifespan. *Nat Neurosci* 18:1413–1420.
- Becker JB, Koob GF (2016) Sex differences in animal models: focus on addiction. *Pharmacol Rev* 68:242–263.
- Becker JB, Prendergast BJ, Liang JW (2016) Female rats are not more variable than male rats: a meta-analysis of neuroscience studies. *Biol Sex Differ* 7:34–37.
- Beltz AM, Beery AK, Becker JB (2019) Analysis of sex differences in pre-clinical and clinical data sets. *Neuropsychopharmacology* 44:2155–2158.
- Beery AK (2018) Inclusion of females does not increase variability in rodent research studies. *Curr Opin Behav Sci* 23:143–149.
- Beery AK, Zucker I (2011) Sex bias in neuroscience and biomedical research. *Neurosci & Biobehav Rev* 35:565–572.
- Berman KF, Schmidt PJ, Rubinow DR, Danaceau MA, Van Horn JD, Esposito G, Ostrem JL, Weinberger DR (1997) Modulation of cognition-specific cortical activity by gonadal steroids: a positron-emission tomography study in women. *Proc Natl Acad Sci USA* 94:8836–8841.
- Bland JM, Altman DG (2015) Best (but oft forgotten) practices: testing for treatment effects in randomized trials by separate analyses of changes from baseline in each group is a misleading approach. *Am J Clin Nutr* 102:991–994.
- Behar MC, Spong C, Grady C, Goldkind SF, Sahin L, Clayton JA (2013) Enrolling pregnant women: issues in clinical research. *Womens Health Issues* 23:e39–e45.
- Bohacek J, Manuella F, Roszkowski M, Mansuy IM (2015) Hippocampal gene expression induced by cold swim stress depends on sex and handling. *Psychoneuroendocrinology* 52:1–12.
- Bolte G, Jacke K, Groth K, Kraus U, Dandolo L, Fiedel L, Debiak M, Kolossa-Gehring M, Schneider A, Palm K (2021) Integrating sex/gender into environmental health research: development of a conceptual framework. *Int J Environ Res Public Health* 18:12118.
- Bond KM, McCarthy MM, Rubin JB, Swanson KR (2021) Molecular omics resources should require sex annotation: a call for action. *Nat Methods* 18:585–588.
- Breedlove SM (1994) Sexual differentiation of the human nervous system. *Annu Rev Psychol* 45:389–418.
- Breedlove SM, Arnold AP (1981) Sexually dimorphic motor nucleus in the rat lumbar spinal cord: response to adult hormone manipulation, absence in androgen-insensitive rats. *Brain Res* 225:297–307.
- Buch T, Moos K, Ferreira FM, Frölich H, Gebhard C, Tresch A (2019) Benefits of a factorial design focusing on inclusion of female and male animals in one experiment. *J Mol Med* 97:871–877.
- Buckley RF, Waller M, Masters CL, Dobson A (2019) To what extent does age at death account for sex differences in rates of mortality from Alzheimer disease? *Am J Epidemiol* 188:1213–1223.
- Carcel C, Harris K, Peters SA, Sandset EC, Balicki G, Bushnell CD, Howard VJ, Reeves MJ, Anderson CS, Kelly PJ, Woodward M (2021) Representation of women in stroke clinical trials: a review of 281 trials involving more than 500,000 participants. *Neurology* 97:e1768–e1774.
- Champagne FA, Curley JP (2009) Epigenetic mechanisms mediating the long-term effects of maternal care on development. *Neurosci Biobehav Rev* 33:593–600.
- Clayton JA (2018) Applying the new SABV (sex as a biological variable) policy to research and clinical care. *Physiol Behav* 187:2–5.
- Clayton JA, Tannenbaum C (2016) Reporting sex, gender, or both in clinical research? *JAMA* 316:1863–1864.
- Cornelison TL, Clayton JA (2017) Article commentary: considering sex as a biological variable in biomedical research. *Gender Genome* 1:89–93.
- De Vries GJ (2004) Minireview: sex differences in adult and developing brains: compensation, compensation, compensation. *Endocrinology* 145:1063–1068.
- Duchesne A, Kaiser Trujillo A (2021) Reflections on neurofeminism and intersectionality using insights from psychology. *Front Hum Neurosci* 15:684412.
- Eliot L (2011) The trouble with sex differences. *Neuron* 72:895–898.
- Eliot L, Ahmed A, Khan H, Patel J (2021) Dump the ‘dimorphism’: comprehensive synthesis of human brain studies reveals few male-female differences beyond size. *Neurosci Biobehav Rev* 125:667–697.
- Fiest KM, Roberts JL, Maxwell CJ, Hogan DB, Smith EE, Frolkis A, Cohen A, Kirk A, Pearson D, Pringsheim T, Venegas-Torres A, Jetté N (2016) The prevalence and incidence of dementia due to Alzheimer’s disease: a systematic review and meta-analysis. *Can J Neurol Sci* 43 Suppl 1:S51–S82.
- Forde NJ, Jeyachandra J, Joseph M, Jacobs GR, Dickie E, Satterthwaite TD, Shinohara RT, Ameis SH, Voineskos AN (2020) Sex differences in variability of brain structure across the lifespan. *Cereb Cortex* 30:5420–5430.
- Galea LA, Choleris E, Albert AY, McCarthy MM, Sohrabji F (2020) The promises and pitfalls of sex difference research. *Front Neuroendocrinol* 56:100817.
- Garcia-Sifuentes Y, Maney DL (2021) Reporting and misreporting of sex differences in the biological sciences. *eLife* 10:e70817.
- Gatewood JD, Wills A, Shetty S, Xu J, Arnold AP, Burgoyne PS, Rissman EF (2006) Sex chromosome complement and gonadal sex influence aggressive and parental behaviors in mice. *J Neurosci* 26:2335–2342.
- Geller SE, Koch AR, Roesch P, Filut A, Hallgren E, Carnes M (2018) The more things change, the more they stay the same: a study to evaluate compliance with inclusion and assessment of women and minorities in randomized controlled trials. *Acad Med* 93:630–635.
- Gelman A, Stern H (2006) The difference between ‘significant’ and ‘not significant’ is not itself statistically significant. *Am Stat* 60:328–331.
- George BJ, Beasley TM, Brown AW, Dawson J, Dimova R, Divers J, Goldsby TU, Heo M, Kaiser KA, Keith SW, Kim MY, Li P, Mehta T, Oakes JM, Skinner A, Stuart E, Allison DB (2016) Common scientific and statistical errors in obesity research. *Obesity* 24:781–790.
- Goerzen D, Fowler C, Devenyi GA, Germann J, Madularu D, Chakravarty MM, Near J (2020) An MRI-derived neuroanatomical atlas of the Fischer 344 rat brain. *Sci Rep* 10:6952.
- Gold EB, Crawford SL, Avis NE, Crandall CJ, Matthews KA, Waetjen LE, Lee JS, Thurston R, Vuga M, Harlow SD (2013) Factors related to age at natural menopause: longitudinal analyses from SWAN. *Am J Epidemiol* 178:70–83.
- Goodfellow PN, Lovell-Badge R (1993) SRY and sex determination in mammals. *Annu Rev Genet* 27:71–92.
- Grant BF, Dawson DA, Stinson FS, Chou SP, Dufour MC, Pickering RP (2004) The 12-month prevalence and trends in DSM-IV alcohol abuse and dependence: United States, 1991–1992 and 2001–2002. *Drug Alcohol Depend* 74:223–234.
- Gruene TM, Flick K, Stefano A, Shea SD, Shansky RM (2015) Sexually divergent expression of active and passive conditioned fear responses in rats. *eLife* 4:e11352.

- Hänggi J, Fövenyi L, Liem F, Meyer M, Jäncke L (2014) The hypothesis of neuronal interconnectivity as a function of brain size: a general organization principle of the human connectome. *Front Hum Neurosci* 8:915.
- Heidari S, Babor TF, De Castro P, Tort S, Curno M (2016) Sex and gender equity in research: rationale for the SAGER guidelines and recommended use. *Res Integr Peer Rev* 1:2–9.
- Hoekzema E, Barba-Müller E, Pozzobon C, Picado M, Lucco F, García-García D, Soliva JC, Tobeña A, Desco M, Crone EA, Ballesteros A, Carmona S, Vilarroya O (2017) Pregnancy leads to long-lasting changes in human brain structure. *Nat Neurosci* 20:287–296.
- Horstmann S, Schmeche C, Palm K, Oertelt-Prigione S, Bolte G (2022) The operationalisation of sex and gender in quantitative health-related research: a scoping review. *Int J Environ Res Public Health* 19:7493.
- Hudson JI, Hiripi E, Pope HG, Kessler RC (2007) The prevalence and correlates of eating disorders in the national comorbidity survey replication. *Biol Psychiatry* 61:348–358.
- Hyde JS (2005) The gender similarities hypothesis. *Am Psychol* 60:581–592.
- Hyde JS, Bigler RS, Joel D, Tate CC, van Anders SM (2019) The future of sex and gender in psychology: five challenges to the gender binary. *Am Psychol* 74:171–193.
- Hyde JS, Mezulis AH (2020) Gender differences in depression: biological, affective, cognitive, and sociocultural factors. *Harv Rev Psychiatry* 28:4–13.
- Ingalhalikar M, Smith A, Parker D, Satterthwaite TD, Elliott MA, Ruparel K, Hakonarson H, Gur RE, Gur RC, Verma R (2014) Sex differences in the structural connectome of the human brain. *Proc Natl Acad Sci USA* 111:823–828.
- Jacobs EG, Weiss BK, Makris N, Whitfield-Gabrieli S, Buka SL, Klibanski A, Goldstein JM (2016) Impact of sex and menopausal status on episodic memory circuitry in early midlife. *J Neurosci* 36:10163–10173.
- Jäncke L (2018) Sex/gender differences in cognition, neurophysiology, and neuroanatomy. *F1000Res* 7:805.
- Jäncke L, Mérillat S, Liem F, Hänggi J (2015) Brain size, sex, and the aging brain. *Hum Brain Mapp* 36:150–169.
- Jain A, Huang GZ, Woolley CS (2019) Latent sex differences in molecular signaling that underlies excitatory synaptic potentiation in the hippocampus. *J Neurosci* 39:1552–1565.
- Kaczurkin AN, Raznahan A, Satterthwaite TD (2019) Sex differences in the developing brain: insights from multimodal neuroimaging. *Neuropsychopharmacology* 44:71–85.
- Klein SL, Schiebinger L, Stefanick ML, Cahill L, Danska J, de Vries GJ, Kibbe MR, McCarthy MM, Mogil JS, Woodruff TK, Zucker I (2015) Sex inclusion in basic research drives discovery. *Proc Natl Acad Sci USA* 112:5257–5258.
- Knoedler JR, Inoue S, Bayless DW, Yang T, Tantry A, Davis CH, Leung NY, Parthasarathy S, Wang G, Alvarado M, Rizvi AH, Fenno LE, Ramakrishnan C, Deisseroth K, Shah NM (2022) A functional cellular framework for sex and estrous cycle-dependent gene expression and behavior. *Cell* 185:654–671.
- Krieger N (2003) Genders, sexes, and health: what are the connections—and why does it matter? *Int J Epidemiol* 32:652–657.
- Labonté B, et al. (2017) Sex-specific transcriptional signatures in human depression. *Nat Med* 23:1102–1111.
- Lewis JD, Theilmann RJ, Sereno MI, Townsend J (2009) The relation between connection length and degree of connectivity in young adults: a DTI analysis. *Cereb Cortex* 19:554–562.
- Levy DR, Hunter N, Lin S, Robinson EM, Gillis W, Conlin EB, Anyoha R, Shansky RM, Datta SR (2023) Mouse spontaneous behavior reflects individual variation rather than estrous state. *Curr Biol* 33:1358–1364. e4.
- Lonsdorf EV, Eberly LE, Pusey AE (2004) Sex differences in learning in chimpanzees. *Nature* 428:715–716.
- Loomes R, Hull L, Mandy WP (2017) What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *J Am Acad Child Adolesc Psychiatry* 56:466–474.
- Makin R, de Xivry JJO (2019) Ten common statistical mistakes to watch out for when writing or reviewing a manuscript. *eLife* 8:e48175.
- Maney DL (2014) Just like a circus: the public consumption of sex differences. *Curr Top Behav Neurosci* 19:279–296.
- Maney DL (2016) Perils and pitfalls of reporting sex differences. *Philos Trans R Soc Lond B Biol Sci* 371:20150119.
- Maney DL, Rich-Edwards JW (2023) Sex-inclusive biomedicine: are new policies increasing rigor and reproducibility? *Womens Health Issues*. Advance online publication. Retrieved Apr 20, 2023. <https://doi.org/10.1016/j.whi.2023.03.004>.
- Maney DL, Karkazis K, Hagen KS (2023) Considering sex as a variable at a research university: knowledge, attitudes, and practice. *J Womens Health*, in press.
- Martínez-García M, Paternina-Die M, Cardenas SI, Vilarroya O, Desco M, Carmona S, Saxbe DE (2023) First-time fathers show longitudinal gray matter cortical volume reductions: evidence from two international samples. *Cereb Cortex* 33:4156–4163.
- Massey SC, Whitmire P, Doyle TE, Ippolito JE, Mrugala MM, Hu LS, Canoll P, Anderson AR, Wilson MA, Fitzpatrick SM, McCarthy MM, Rubin JB, Swanson KR (2021) Sex differences in health and disease: a review of biological sex differences relevant to cancer with a spotlight on glioma. *Cancer Lett* 498:178–187.
- McCarthy MM (2016) Multifaceted origins of sex differences in the brain. *Philos Trans R Soc Lond B Biol Sci* 371:20150106.
- McCarthy MM (2020) A new view of sexual differentiation of mammalian brain. *J Comp Physiol A* 206:369–378.
- McCarthy MM, Arnold AP (2011) Reframing sexual differentiation of the brain. *Nat Neurosci* 14:677–683.
- McCarthy MM, Konkle AT (2005) When is a sex difference not a sex difference? *Front Neuroendocrinol* 26:85–102.
- McCarthy MM, Arnold AP, Ball GF, Blaustein JD, De Vries GJ (2012) Sex differences in the brain: the not so inconvenient truth. *J Neurosci* 32:2241–2247.
- McCarthy MM, Herold K, Stockman SL (2018) Fast, furious and enduring: sensitive versus critical periods in sexual differentiation of the brain. *Physiol Behav* 187:13–19.
- McLean CP, Asnaani A, Litz BT, Hofmann SG (2011) Gender differences in anxiety disorders: prevalence, course of illness, comorbidity and burden of illness. *J Psychiatr Res* 45:1027–1035.
- Miani C, Wandtschneider L, Niemann J, Batram-Zantvoort S, Razum O (2021) Measurement of gender as a social determinant of health in epidemiology: a scoping review. *PLoS One* 16:e0259223.
- Mielke MM, Ferretti MT, Iulita MF, Hayden K, Khachaturian AS (2018) Sex and gender in Alzheimer's disease: does it matter? *Alzheimers Dement* 14:1101–1103.
- Moore CL (1992) The role of maternal stimulation in the development of sexual behavior and its neural basis. *Ann NY Acad Sci* 662:160–177.
- National Academies of Sciences, Engineering, and Medicine (2022) Measuring sex, gender identity, and sexual orientation. Washington, DC: National Academies Press.
- National Academies of Sciences, Engineering, and Medicine (2023) Using population descriptors in genetics and genomics research: a new framework for an evolving field. Washington, DC: National Academies Press.
- National Institutes of Health (2015) Reviewer guidance to evaluate sex as a biological variable (SABV). Available at https://grants.nih.gov/grants/peer/guidelines_general/SABV_Decision_Tree_for_Reviewers.pdf. Accessed August 5, 2023.
- National Institutes of Health (2020) Sex as a biological variable: a primer. Available at <https://orwh.od.nih.gov/career-development-education/e-learning/sabv-primer>. Accessed August 5, 2023.
- Nielsen MW, Stefanick ML, Peragine D, Neilands TB, Ioannidis JP, Pilote L, Prochaska JJ, Cullen MR, Einstein G, Klinge I, LeBlanc H, Paik HY, Schiebinger L (2021) Gender-related variables for health research. *Biol Sex Differ* 12:23–16.
- Nieuwenhuis S, Forstmann BU, Wagenmakers EJ (2011) Erroneous analyses of interactions in neuroscience: a problem of significance. *Nat Neurosci* 14:1105–1107.
- Nottebohm F, Arnold AP (1976) Sexual dimorphism in vocal control areas of the songbird brain. *Science* 194:211–213.
- O'Connor C, Joffe H (2014) Gender on the brain: a case study of science communication in the new media environment. *PLoS One* 9:e110830.
- Özdemir BC, Gerard CL, Espinosa da Silva C (2022) Sex and gender differences in anticancer treatment toxicity: a call for revisiting drug dosing in oncology. *Endocrinology* 163:bqac058.
- Pelletier R, Khan NA, Cox J, Daskalopoulou SS, Eisenberg MJ, Bacon SL, Lavoie KL, Daskupta K, Rabi D, Humphries KH, Norris CM, Thanassoulis G, Behloul H, Pilote L, GENESIS-PRAXY Investigators (2016) Sex versus gender-related characteristics: which predicts outcome after acute coronary syndrome in the young? *J Am Coll Cardiol* 67:127–135.

- Petersen N, Beltz AM, Casto KV, Taylor CM, Jacobs EG, Sundström-Poromaa I, Pletzer B (2023) Towards a more comprehensive neuroscience of hormonal contraception. *Nat Neurosci* 26:529–531.
- Phillips B, Haschler TN, Karp NA (2023) Statistical simulations show that scientists need not increase overall sample size by default when including both sexes in in vivo studies. *PLoS Biol* 21:e3002129.
- Pintzka CW, Hansen TI, Evensmoen HR, Håberg AK (2015) Marked effects of intracranial volume correction methods on sex differences in neuroanatomical structures: a HUNT MRI study. *Front Neurosci* 9:238.
- Polanczyk G, De Lima MS, Horta BL, Biederman J, Rohde LA (2007) The worldwide prevalence of ADHD: a systematic review and meta-regression analysis. *Am J Psychiatry* 164:942–948.
- Prendergast BJ, Onishi KG, Zucker I (2014) Female mice liberated for inclusion in neuroscience and biomedical research. *Neurosci Biobehav Rev* 40:1–5.
- Pritschet L (2022) The scientific body of knowledge: whose body does it serve? A spotlight on human neuroscience. National Institutes of Health Science Policy Scholar Talk presented at the Organization for the Study of Sex Differences, Marina Del Ray, CA.
- Pritschet L, Santander T, Taylor CM, Layher E, Yu S, Miller MB, Grafton ST, Jacobs EG (2020) Functional reorganization of brain networks across the human menstrual cycle. *Neuroimage* 220:117091.
- Pritschet L, Taylor C, Santander T, Jacobs EG (2021) Applying dense-sampling methods to reveal dynamic endocrine modulation of the nervous system. *Curr Opin Behav Sci* 40:72–78.
- Raam T, Hong W (2021) Organization of neural circuits underlying social behavior: a consideration of the medial amygdala. *Curr Opin Neurobiol* 68:124–136.
- Rauch JM, Eliot L (2022) Breaking the binary: gender versus sex analysis in human brain imaging. *Neuroimage* 264:119732.
- Rechlin RK, Splinter TF, Hodges TE, Albert AY, Galea LA (2022) An analysis of neuroscience and psychiatry papers published from 2009 and 2019 outlines opportunities for increasing discovery of sex differences. *Nat Commun* 13:2137.
- Richardson SS (2022) Sex contextualism. *Philos Theory Practice Biol* 14:100550.
- Rippon G, Eliot L, Genon S, Joel D (2021) How hype and hyperbole distort the neuroscience of sex differences. *PLoS Biol* 19:e3001253.
- Ritchie SJ, Cox SR, Shen X, Lombardo MV, Reus LM, Alloza C, Harris MA, Alderson HL, Hunter S, Neilson E, Liewald DC, Auyeung B, Whalley HC, Lawrie SM, Gale CR, Bastin ME, McIntosh AM, Deary IJ (2018) Sex differences in the adult human brain: evidence from 5216 UK Biobank participants. *Cereb Cortex* 28:2959–2975.
- Rosenthal R (1979) The file drawer problem and tolerance for null results. *Psychol Bull* 86:638–641.
- Rutter M, Caspi A, Fergusson D, Horwood LJ, Goodman R, Maughan B, Moffitt TE, Meltzer H, Carroll J (2004) Sex differences in developmental reading disability: new findings from 4 epidemiological studies. *JAMA* 291:2007–2012.
- Ryan BC, Vandenberg JG (2002) Intrauterine position effects. *Neurosci Biobehav Rev* 26:665–678.
- Sainani K (2010) Misleading comparisons: the fallacy of comparing statistical significance. *PM & R* 2:559–562.
- Saini A (2017) *Inferior: how science got women wrong*. Boston: Beacon.
- Salk RH, Hyde JS, Abramson LY (2017) Gender differences in depression in representative national samples: meta-analyses of diagnoses and symptoms. *Psychol Bull* 143:783–822.
- Schiebinger L (1993) *Nature's body: gender in the making of modern science*. Boston: Beacon.
- Shors TJ, Chua C, Falduto J (2001) Sex differences and opposite effects of stress on dendritic spine density in the male versus female hippocampus. *J Neurosci* 21:6292–6297.
- Smarr BL, Ishami AL, Schirmer AE (2021) Lower variability in female students than male students at multiple timescales supports the use of sex as a biological variable in human studies. *Biol Sex Differ* 12:32.
- Smith-Bouvier DL, Divekar AA, Sasidhar M, Du S, Tiwari-Woodruff SK, King JK, Arnold AP, Singh RR, Voskuhl RR (2008) A role for sex chromosome complement in the female bias in autoimmune disease. *J Exp Med* 205:1099–1108.
- Sorge RE, Strath LJ (2018) Sex differences in pain responses. *Curr Opin Physiol* 6:75–81.
- Spring S, Lerch JP, Henkelman RM (2007) Sexual dimorphism revealed in the structure of the mouse brain using three-dimensional magnetic resonance imaging. *Neuroimage* 35:1424–1433.
- Springer KW, Mager Stellman J, Jordan-Young RM (2012) Beyond a catalogue of differences: a theoretical frame and good practice guidelines for researching sex/gender in human health. *Soc Sci Med* 74:1817–1824.
- Sudai M, Borsari A, Ichikawa K, Shattuck-Heidorn H, Zhao H, Richardson SS (2022) Law, policy, biology, and sex: critical issues for researchers. *Science* 376:802–804.
- Tabatadze N, Huang G, May RM, Jain A, Woolley CS (2015) Sex differences in molecular signaling at inhibitory synapses in the hippocampus. *J Neurosci* 35:11252–11265.
- Tapp A, Mayberry MT, Whitehouse AJ (2011) Evaluating the twin testosterone transfer hypothesis: a review of the empirical evidence. *Horm Behav* 60:713–722.
- Taylor C, Pritschet L, Yu S, Jacobs EG (2019) Applying a women's health lens to the study of the aging brain. *Front Hum Neurosci* 13:224.
- Taylor C, Pritschet L, Olsen R, Layher E, Santander T, Grafton ST, Jacobs EG (2020) Progesterone shapes medial temporal lobe volume across the human menstrual cycle. *Neuroimage* 220:117125.
- Taylor C, Pritschet L, Jacobs EG (2021) The scientific body of knowledge: whose body does it serve? A spotlight on oral contraceptives and women's health factors in neuroimaging. *Front Neuroendocrinol* 60:100874.
- U.S. Department of Health and Human Services (2023) Sex & gender. <https://orwh.od.nih.gov/sex-gender>. Accessed August 5, 2023.
- Vijayakumar N, de Macks ZO, Shirtcliff EA, Pfeifer JH (2018) Puberty and the human brain: insights into adolescent development. *Neurosci Biobehav Rev* 92:417–436.
- Vorland CJ (2021) Sex difference analyses under scrutiny. *eLife* 10:e74135.
- Vorland CJ, Brown AW, Dawson JA, Dickinson SL, Golzarri-Arroyo L, Hannon BA, Heo M, Heymsfield SB, Jayawardene WP, Kahathuduwa CN, Keith SW, Oakes JM, Tekwe CD, Thabane L, Allison DB (2021) Errors in the implementation, analysis, and reporting of randomization within obesity and nutrition research: a guide to their avoidance. *Int J Obes* 45:2335–2346.
- Vorland CJ, Golzarri-Arroyo L, Maney DL, Brown AW (2023) Determination of sex differences requires formal test for differences: comment on 'Lipoprotein-subclass particle numbers in children with abdominal obesity.' *Pediatr Int* 65:e15489.
- Wallen K, Baum MJ (2002) Masculinization and defeminization in altricial and precocial mammals: comparative aspects of steroid hormone action. In: *Hormones, brain and behavior* (Pfaff D, ed), pp 385–424. London: Academic.
- White J, Tannenbaum C, Klinge I, Schiebinger L, Clayton J (2021) The integration of sex and gender considerations into biomedical research: lessons from international funding agencies. *J Clin Endocrinol Metab* 106:3034–3048.
- Wierenga LM, et al., Karolinska Schizophrenia Project (KaSP) Consortium (2022) Greater male than female variability in regional brain structure across the lifespan. *Hum Brain Mapp* 43:470–499.
- Williams CM, Peyre H, Toro R, Ramus F (2021) Neuroanatomical norms in the UK Biobank: the impact of allometric scaling, sex, and age. *Hum Brain Mapp* 42:4623–4642.
- Wilson LA, Zajitschek SR, Lagisz M, Mason J, Haselimahhadi H, Nakagawa S (2022) Sex differences in allometry for phenotypic traits in mice indicate that females are not scaled males. *Nat Commun* 13:7502.
- Woitovich NC, Beery A, Woodruff T (2020) Meta-research: a 10-year follow-up study of sex inclusion in the biological sciences. *eLife* 9:e56344.
- Woolley CS, Gould E, Frankfurt M, McEwen BS (1990) Naturally occurring fluctuation in dendritic spine density on adult hippocampal pyramidal neurons. *J Neurosci* 10:4035–4039.
- Zell E, Krizan Z, Teeter SR (2015) Evaluating gender similarities and differences using metasynthesis. *Am Psychol* 70:10–20.
- Zhang K, Sejnowski TJ (2000) A universal scaling law between gray matter and white matter of cerebral cortex. *Proc Natl Acad Sci USA* 97:5621–5626.
- Zsido R, Williams A, Barth C, Serio B, Kurth L, Beyer F, Witte AV, Villringer A, Sacher J (2023) Menstrual cycle brain plasticity: ultra-high field 7T MRI reveals changes in human medial temporal lobe volume in female adults. *Nat Hum Behav*, in press.